数学小丛书

18

费马猜想

冯克勤

科学出版社
www.sciencep.com

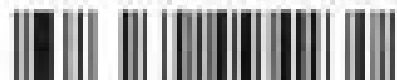
(O-1389.0101) 责任编辑：吕 虹/总体设计：黄华斌/责任印制：安春生

18

数学小丛书

O1-
H68

ISBN 7-03-009423-9



数学小丛书 18

费马猜想

冯 克 勤

科 学 出 版 社
2002

内 容 简 介

1637年法国数学家费马提出一个数学猜想，于1994年由怀尔斯给出证明，被认为是20世纪纯粹数学的一项重大成就.证明中使用了近年来在代数、数论和几何学方面的许多重大研究成果.本书较为通俗地介绍300多年来人们攻克费马猜想的历史进程，在解决费马猜想中产生的创新思想和方法，以及对发展数学的推进作用.

图书在版编目(CIP)数据

费马猜想/冯克勤.—北京:科学出版社,2002
(数学小丛书)
ISBN 7-03-009423-9

Ⅰ.费… Ⅱ.冯… Ⅲ.费马最后定理-普及读物
Ⅳ.O156-49

中国版本图书馆CIP数据核字(2002)第010499号

科学出版社 出版
北京东黄城根北街16号
邮政编码:100717
http://www.sciencep.com

中国科学院印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2002年5月第一版 开本:787×960 1/32
2002年5月第一次印刷 印张:6 3/8 插页:1
印数:1—5 000 字数:97 000

全套书定价:99.00元 (共18册)

馬克思说：「一门科学，只有当它成功地运用数学时，才能达到真正完善的地步。」恩格斯说：「要辩证而又唯物地了解自然，就必须熟悉数学。」在科教兴国，振兴中华的今天，向全社会普及数学，实在是一件刻不容缓的大事。

数学小丛书是由我国一些著名数学家撰写的一批数学普及读物精品。几十年来，我国几代科技人员中，不少人都曾得益于这套丛书。我衷心地祝贺数学小丛书的重版与补充，并预祝它取得更大的成功。

王元 

二〇〇〇年四月

21

数学天元基金

出版说明

1956年，为了向青少年传播数学知识，科学出版社配合我国首次举办的高中数学竞赛，出版了老一辈数学家华罗庚教授的《从杨辉三角谈起》和段学复教授的《对称》. 在20世纪60年代初，这两本书连同其他一些著名数学家撰写的科普著作，被北京市数学会编成小丛书，相继由不同的出版社出版，并多次重印.

由数学大师和著名数学家亲自执笔撰写的这套数学小丛书是我国数学普及读物中的精品，曾激发一代青少年学习数学的兴趣. 书中蕴涵的深刻而富有启发性的思想，促进了无数中学生在求学的道路上健康成长. 当年这套小丛书的许多读者，现在已经成为学有所成的科学技术工作者，国家建设的栋梁之才. 当年由老一辈数学家所倡导的我国的数学竞赛活动，现在已经得到蓬勃的发展. 我国自1986年正式参加国际数学奥林匹克竞赛以来，历届都取得总分

第一或第二的好成绩.近年来,我国的数学普及读物无论是品种还是数量都在增加,但是这套数学小丛书仍然无愧是其中别具特色的瑰宝,理应成为传世之作.因此,我社取得作者或其继承人的同意,并在可能的条件下,请作者本人或相关学者对重新编辑的书稿进行了审订,重新刊行这套数学小丛书,以飨广大青少年读者.

数学是几千年人类智慧的结晶,是一门古老而又常新的科学.借此丛书再版之机,我们特别增加两本新书:虞言林教授等的《祖冲之算 π 之谜》和冯克勤教授的《费马猜想》.前者介绍中国古代数学的一项重大成就,后者阐述数学史上的一个著名猜想——费马定理历经 300 多年终于在 20 世纪末被证明的故事,我们相信读者从中将会受到启迪.

本套丛书以新貌重新出版,得到了国家自然科学基金委员会数学天元基金的资助,谨表示衷心感谢.

前　　言

1637 年,法国数学家费马(Fermat)提出如下的猜想:对于每个大于 2 的正整数 n,任意两个正整数的 n 次方之和不能为另一个正整数的 n 次方.也就是说,方程 $x^n + y^n = z^n$ 没有正整数解(x, y, z).这就是著名的费马猜想.这个猜想是如此简单易懂,可是要证明却出人意料地艰难.300 多年来,许多专业数学家和业余数学爱好者为解决此猜想作了不懈的努力,其中包括像欧拉(Euler)等一些大数学家的努力,最终于 1994 年由 41 岁的英国数学家怀尔斯(A. Wiles)所证明.这项工作被认为是本世纪最重要的理论数学成就之一.

费马猜想是关于整数性质的一个论断,它属于数学的一个古老分支:数论,这是研究整数性质和方程整数解的一种学问.数论中有许多简单易懂的问题和猜想,其中有一些至今仍未解决(例如哥德巴赫问题).这些数论难题是对人类智慧的一种挑战,而人们为解决这些难题所作的贡献,是对数论乃至整个数学发展的巨

大推动．在费马猜想的研究过程中，创造了研究数论的许多新方法，建立了数论的新分支，使用了几何、代数和分析学的各种工具，发现了费马猜想与数学其他领域的奇妙而深刻的联系．它的意义远远超过了证明费马猜想这个事实本身．费马猜想的证明历程充分显示出，现今蓬勃发展的数学是一个有机的整体，不同数学分支的相互交叉和渗透是产生数学新思想和创造数学重大成果的源泉．

高斯称数论是“数学的皇后”．当今的数论已发展成一门十分艰深的学问，例如怀尔斯对费马猜想的证明就是非常难懂的．但是近年来，数论在实际领域中得到广泛而深刻的应用．这主要是由于 20 世纪 60 年代以来计算机技术和数字通信技术的飞速发展，使得数论和其他离散数学成为计算机科学和通信工程的重要数学工具．特别是一些重大数论成果的应用，使技术领域发生了巨大变革，显示了纯粹数学理论对实际的推动作用．

本书向读者展示费马猜想的历史进程．基于本书的通俗性要求和限于作者的能力，我们不可能讲述费马猜想的严格证明，但试图勾画人们在研究费马猜想过程中所创造和使用的数学思想和方法，介绍与费马猜想有关联的近代数论的发展．

冯 克 勤

1999 年 1 月

目　　录

1 数(shù)起源于数(shǔ)

我们要用五小节的篇幅来讲述费马猜想之前的数论,介绍数论是如何产生的,在费马猜想之前人类所掌握的数论知识.在介绍费马猜想之前的数论历史的过程中,我们将着重讲述与费马猜想有关的初等数论重要内容、概念、术语和符号.这一节先谈谈数论的起源和古代数论的发展.

和艺术、天文一样,数学是人类最古老的精神文明之一,数学的萌芽产生于文字发明之前,距今至少有六七千年的历史.古代人类聚居在气候温和、空气湿润、土壤肥沃的大河流域.这就是尼罗河流域的埃及,底格里斯河和幼发拉底河流域的巴比伦(现在伊拉克的地方),恒河流域的印度和黄河长江流域的中国.数学起源于这四大文明古国.

数(shù)起源于数(shǔ),量(liàng)起源于量(liáng).人有生老病死,每个氏族部落的成员经

常发生变化(增多或减少);每次狩猎归来,需要估量猎物的多寡,分配食物时需要把猎物和氏族成员的多少加以比较;尼罗河每年洪水泛滥,洪水退去之后,需要重新丈量土地;建筑房屋、堤坝和巨大的金字塔,需要计算各种图形的面积和体积,所以数学产生于对数量的认识和对几何图形的认识,而最早认识到的数是 1,2,3 这些正整数.

现在,每个幼儿园的孩子都可以数出 1,2,3 及更大的数字,但是在几千年之前,人们从 3 只羊、3 个人和 3 块石头中间提炼出它们共同的性质,产生了数 3 的概念,是非常不简单的.考古学家发现,在有文字之前,人们是用石子、沙粒、树枝和贝壳等实物来计数的.1930 年,美国的考古队在伊拉克境内发现一个封口泥罐,泥罐表面画着一种牲畜,罐里有 48 颗泥粒,这表示泥罐的主人曾经有过 48 头这种牲畜.中国史书上有“上古结绳而治”一说,人们在绳上打几个结,用来记载有几个事物.对于少量物品,人们用手指计数,物品多了则用树枝在泥巴上刻痕,或用刀具在动物骨骼上刻线.

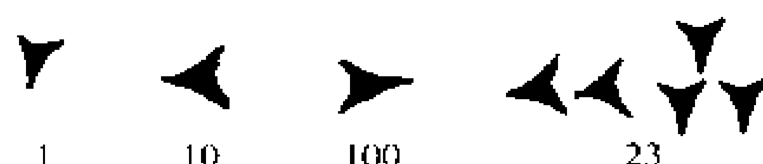
大约在公元前 4000 年,人类发明了文字.各种数目以固定的形式书写成文字的形式,这就是数字.看一看各文明古国不同的数字表达方式是非常有趣的.在埃及,最初的文字是象形

的，用树枝蘸着炭汁，写在芦苇挤压晒干而成的纸草上，这些数字为：

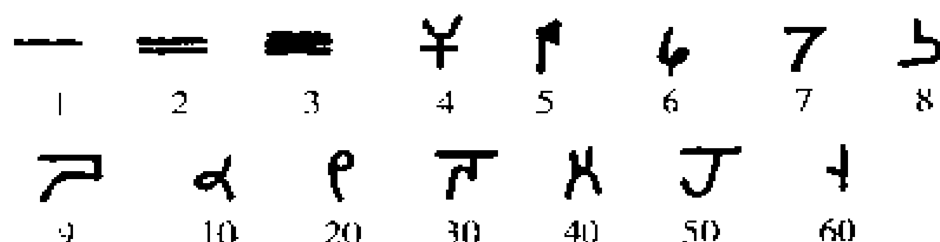


例如，6 表示成 $\substack{|||\\|||}$，300 表示成 等等.

在巴比伦，用削尖的木棒在半湿的软泥板上书写文字，每个笔画的形状是楔形，称作楔形文字，数字写成：



在印度，公元前 2 世纪数字表示成如下的婆罗门式记号：



大家会发现，其中的 6 和 7 与现在表达方式很相像.事实上，在 13 世纪初，阿拉伯人把印度数字加以变化传到欧洲，被欧洲人称为阿拉伯数字，就是我们今天普遍采用的数字.

在中国，数目字也出现得很早，距今约

6000 年前的西安半坡村新石器时代遗址中有刻在陶器上的数字：

X (5)　∧ (6)　十 (7)　)((8)　| (10)　|| (20)

我国系统的数目字大约出现在商代，用甲骨文书写的数字有：

一	二	三	亖	X	∧ 或 ∧	十	)(				
1	2	3	4	5	6	7	8	9	10

另有百、千、万等高位值符号：

（百）或（百）　（千）　（万）

到了春秋时期（公元前 700 ~ 公元前 476 年），我们的祖先创造了用算筹表示数字和进行运算的“筹算”. 算筹通常用竹子刻成，形如筷子，汉朝的算筹长约 13 厘米，用算筹摆出的数字有纵横两种形式：

	1	2	3	4	5	6	7	8	9
纵式:	𝍠	𝍡	𝍢	𝍣	𝍤	𝍥	𝍦	𝍧	𝍨
横式:	𝍩	𝍪	𝍫	𝍬	𝍭	𝍮	𝍯	𝍰	𝍱

记数时个位常用纵式，依次横纵相间，如遇零便空一位. 如 6143 的筹式为 𝍮𝍠𝍬𝍢，

306 的筹式为 𝍢　𝍥．这种计数方法实质上就是现在采用的十进位记数方法．在明代珠算盘被普遍使用之前，我国古代一直用算筹进行四则运算．春秋时期中国就有了计算乘法的口诀：九九表（一一得一，一二得二，二二得四，……一直到九九八十一）．不仅如此，中国也是最早认识到分数并且建立分数运算的国家．大约在战国末期，我国数学家就把分数看成是两数相除，用算筹表示是被除数放在除数的上面，例如$\frac{16}{5}$表示成：

𝍩𝍥
𝍤

这与现在的分数记法不同之处只是差一条分数线．在我国著名的古算书《九章算术》（写于公元前 1 世纪）中已经有通分、约分及分数四则运算等相当完整的分数理论，比当时埃及等其他国家要先进很多．

随着生产的发展和生活的进步，如何表示大的数目，是一个非常重要和严肃的问题．开始时，一些大的数目用专门的文字或符号表示，比如前面所示，埃及人把 100 记为 𓍢，把 10 万记成像小鸟一样，中国的甲骨文也把 100，1000，10000 表示成特定的符号．这种方法在某种程

度上可以记下大数，可是运算很不方便.

新记数方法的发明和普遍采用，是古代数学的一个重大进步，这就是发明了“定位制”.以我们现在采用的十进制为例，我们只需要十个符号0,1,2,3,4,5,6,7,8,9就可以表示任意大的正整数.每个数从右到左依次为“个位”，“十位”，“百位”，等等.例如213的3在个位表示3，而1在十位表示10，数字2在百位表示200，所以213为200+10+3，表示贰佰壹拾叁这个数.也就是说，每个数字代表的数值由它所处的位置决定.两个数相加时，相同位置上的数字相加，超过10时则向前(即向左)进位，减法则需要“借位”.这种定位制表示数的方法和运算方法非常方便，一直使用到今天.中国的筹算是世界上最早的十进制计数和运算的工具，后人的改进只是采用了更方便的阿拉伯数字符号.

在东方文明国家采用定位制计数和运算之后，欧洲一直保留陈旧的记数方式：古罗马数字，这种数字一直到今天还用在时钟钟面、日历、书稿的章节分类等方面，大约在公元前后罗马人用7个基本符号表示数：

I	V	X	L	C	D	M
1	5	10	50	100	500	1000

例如3表示成III，而800就要连写8个C或者

写成 DCCC，后来为了简化又创造了一个新规则，即数值较小的符号放在数值较大符号之左边时，则从大数值减去小数值，例如 VI 表示 6，而 IV 表示 4（不写成 IIII）.同样，CD 是 400 而 DC 是 600，于是 89 可以简写成 XXCIX.在罗马帝国灭亡（476 年）后的 700 多年间，西欧人仍然使用这种过于复杂的罗马数字，这也是造成在相当长的一段时间里，西方数学落后于东方数学的原因之一.从这段历史可以看出，一套好的数学符号对于数学的发展是多么重要！

有了整数概念、整数表达方式和方便的记数方法，由于生产和生活的需要发明了整数的四则运算，并且要解决关于整数的各种实际问题，这就产生了研究整数性质和方程整数解的学问：数论.数论的历史大约有 3000 年，起源于古代的东方.中国最早的数学著作《周髀算经》（大约写于公元前 235 年至公元前 145 年之间）的开篇就记载了西周人商高知道方程 $x^2+y^2=z^2$ 有整数解 $(x,y,z)=(3,4,5)$.另一部数学著作《孙子算经》（公元 4～5 世纪）载有"物不知数"问题，研究整数的同余性质，被世人称为"中国剩余定理".东方各国的数论（乃至整个数学）主要基于实践，具有鲜明的直观、实用和算法特性.而在古希腊数学那里（公元前 6 世纪至公元 3 世纪），数论（和几何）的研究具有理性思

辨的特点，古希腊数学成为世界数学史的一个辉煌时期，这有着深刻的地理因素和社会根源.

公元前 8 世纪到公元前 6 世纪，希腊人从半农半牧的氏族公社一跃而为以手工业和商业为主的奴隶制城邦国家的联合体，从地中海东北角落的僻居乡民发展成精通航海贸易，控制了整个地中海航道的当时欧洲最先进民族.这种社会发展过程促使希腊人不断革新、破除保守和勇于探索未知领域.他们从近邻埃及和巴比伦那里学习了历史悠久的文化遗产.他们有充分发展的奴隶制社会分工，使一部分贵族和自由民有足够多时间用于学习和研究.他们崇尚思辨，追求以演绎推导的方式得出普遍适用的一般真理.

在古代大多数民族中，人们对于自然充满恐惧的心理.在雷鸣电闪面前，在风暴和洪水中为生活和生存而挣扎.人在自然面前感到软弱无力，把各种自然现象归结于神的意旨.但是希腊人不仅在身体上直立起来，而且在精神上也站了起来.他们认为各种自然现象是有规律的，而且人是可以认识自然现象的规律的.人们用什么来认识自然规律呢？他们的答案是用数学.人们认识世界的一个最根本问题是：世界是由什么组成的？这是古代各民族哲学家的思考对象.在中国曾有“五行说”，即世界是由“金，

木，水，火，土"构成的．在古希腊哲学家中最早致力于探讨万物本原的是哲学家和数学家泰勒斯（约公元前 625～公元前 547），但影响最大的是他的学生毕达哥拉斯（约公元前 580～公元前 500），他认为万物之本原是数，即所谓"万物皆数"．这一学派认为："任何一种东西之所以能够被认识，是因为包含一种数；没有这种数，心灵什么东西也不能思考，什么东西也不能认识．"正是这种哲学思想促使希腊人对于整数的研究到了入迷的地步，使他们对于整数抽象性质的关心超过了对世俗生活的需要．所以，在古希腊哲学和数学是在一起的．数学是哲学的最重要思考方式和手段，在社会中具有极为重要的地位．

古希腊数学的最重要数论成就集中反映在两本数学著作中．一本是欧几里得（公元前 330 年至公元前 275 年）的著作《几何原本》．这本著作共 13 卷，其中有 3 卷讲述数论．现在列举书中最重要的几项成果．首先，书中讲述了初等数论的基石：

算术基本定理　每个大于 1 的整数均可惟一地表示成有限个素数的乘积．

这个定理给出了正整数进行（乘性）分解的重要性质．一个大于 1 的正整数 n 称为素数（或称为质数），是指它不能写成两个均小于 n

的正整数的乘积(或者说,除了1和 n 之外,n 没有其他正整数因子).如果 $n(\geq 2)$ 不是素数,则 n 可表示成 ab,其中 a 和 b 均是比 n 小的正整数(所以 a 和 b 也都大于 1).如果 a 和 b 仍不是素数,再继续分解,由于分解后的因子愈来愈小,所以经过有限步分解之后,每个因于都是素数.因此 n 一定能表示成有限个素数的乘积:$n=p_1p_2\cdots p_t$(今后我们常用 p 表示素数,多个素数表示成 $p_1,p_2,\cdots$)定理的更重要部分是说将 n 表示成素数乘积的表达方式(本质上)是惟一的.确切地说,如果 $n(\geq 2)$ 可以用两种方式表达成素数乘积:

$$n=p_1p_2\cdots p_t=p'_1p'_2\cdots p'_s,$$

则必然 $t=s$,即是同样多个(t 个)素数的乘积,并且素数 $p_1,\cdots,p_t$ 和素数 $p'_1,\cdots,p'_t$ 至多只是前后次序的不同,若不考虑次序,则它们是同样的 t 个素数.这是相当深刻的结果.古希腊人给出了相当严密的证明.我们将会看到,在费马猜想的早期研究中,主要是得益于以更高的视野对这个惟一因子分解定理重新加以审视.

作正整数分解时,素数是最基本的"原子",所有大于1的正整数都是由素数相乘得来.希腊人问:素数有多少?他们的答案是:素数有无穷多个.《几何原本》中给出了证明,这个证明现

在许多中学生都知道：假如素数只有有限多个：$p_1,\cdots,p_t$.让我们考虑正整数 $n=p_1p_2\cdots p_t+1$.按着算术基本定理，n 应当是一些素数的乘积.但已假定素数只有 $p_1,p_2,\cdots,p_t$，所以这 t 个素数当中必有某一个为 n 的因子.另一方面，$p_1,\cdots,p_t$ 均不能除尽 $n=p_1\cdots p_t+1$，即均不能是 n 的因子，导致矛盾.这就证明了素数必有无穷多个.这可能是数学历史上第一个反证法的证明.除了这种纯思辨的结果之外，古希腊人还对任意正整数 n 给出一种实际方法（"筛法"），求出不超过 n 的所有素数.100 以内的素数为：2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71,73,79,83,89,97.

第三个重要结果是欧几里得算式，即通常所说的"带余除式"：用正整数 b 去除正整数 a，惟一地得到商 q 和余数 r，即

$$a=bq+r,$$

其中 q 是非负整数，余数 r 为整数，并且 $0\leqslant r<b$.利用带余除式，欧几里得还给出求两个整数最大公因子的辗转相除法.

我们要介绍的最后一个结果是：欧几里得的书中给出了方程 $x^2+y^2=z^2$ 的全部正整数解.要注意此方程的正整数解有无穷多个.书中利用整数的性质推导出所有正整数解的表达

式.我们将在下节讲这件事.2000 多年前古希腊的这些数论结果,在今天看来仍是令人惊叹的.

古代数学中对数论的研究常常与几何学结合在一起.例如将方程 $x^2+y^2=z^2$ 的解 x,y,z 看成是直角三角形三条边的长度,在中国称之为勾、股和弦.古希腊的另一重要数学著作是丢番图的《算术》(公元 3 世纪),这是历史上第一部脱离开几何学完全讲述数论的著作.这本书研究了 300 多个数论问题,列举了一些一次和二次方程(组)的有理数解和整数解的各种方法.这本书对于费马猜想具有特殊的意义,因为正是费马在 1000 多年之后阅读此书时提出了费马猜想!

以上我们粗略地介绍了古代数论(或叫初等数论)的历史,下面四节我们就初等数论的一些重要内容作稍微仔细地叙述.

2 算术基本定理

每个大于 1 的整数 n 均可表示成有限个素数的乘积：$n = p_1 p_2 \cdots p_t$，并且若不考虑这些素因子的次序，这个表达式是惟一的.所以若把素因子从小到大排列：$p_1 \leqslant p_2 \leqslant \cdots \leqslant p_t$，那么上面的分解式就完全是惟一的.如果再把相同素数因子的乘积写成方幂形式（如 $100 = 2 \times 2 \times 5 \times 5 = 2^2 \times 5^2$），那么算术基本定理就可表达成如下形式：

每个大于 1 的正整数 n 均可惟一地分解成

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}, \tag{2.1}$$

其中 $p_1, p_2, \cdots, p_s$ 是素数，$p_1 < p_2 < \cdots < p_s$，而 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 均是正整数.公式(2.1)称为 n 的**标准分解式**.

我们说这条定理是数论的基石，是因为用它可研究整数的许多性质.整数的一个重要性

质是整除性问题.首先我们介绍一些术语和符号,其中大多数术语是大家所熟悉的.

设 a 和 b 是整数,其中 $a\neq 0$.如果存在整数 c,使得 $b=ac$,我们称 a 整除 b(或称 b 被 a 整除),表示成 $a\mid b$.而称 a(和 c)为 b 的因子,称 b 为 a(和 c)的倍数.换句话说,a 整除 b 是指 b/a 为整数.如果 a 不整除 b(即 b/a 不是整数),则表示成 $a\nmid b$.例如,$(-3)\mid 6$,$3\mid(-6)$,$4\nmid 6$.对于每个整数 n,$(\pm 1)\mid n$.而对于每个非零整数 n,$n\mid 0$.

若 c 是整数 a 和 b 的因子,则 c 称为 a 和 b 的公因子.若 c 同时是 a 和 b 的倍数,则 c 称为 a 和 b 的公倍数.如果整数 a 和 b 不全为 0,则 a 和 b 必然存在最大公因子,表示成 (a,b).例如,$(4,6)=2$,$(-5,0)=5$.若 a 和 b 均是不为零的整数,则一定存在最小的正公倍数,称为 a 和 b 的最小公倍数,表示成 $[a,b]$.例如,$[-2,-3]=[2,3]=6$,$[4,-6]=12$.类似的,可以定义多个整数 $a_1,a_2,\cdots,a_n$ 的最大公因子和最小公倍数,它们分别表示成 $(a_1,a_2,\cdots,a_n)$ 和 $[a_1,a_2,\cdots,a_n]$.

利用正整数 M 和 N 的素因子分解式,我们容易判别何时 M 整除 N,也容易求出它们的最大公因子和最小公倍数,设

$$M=p_1^{a_1}p_2^{a_2}\cdots p_s^{a_s},\quad N=p_1^{b_1}p_2^{b_2}\cdots p_s^{b_s},\quad (*)$$

其中 $p_1, p_2, \cdots, p_s$ 是不同的素数，而 $a_1, \cdots, a_s, b_1, \cdots, b_s$ 是非负整数. 如果 M 整除 N，则有整数 L 使得 $N = ML$. 所以

$$p_1^{b_1} p_2^{b_2} \cdots p_s^{b_s} = N = ML = p_1^{a_1} p_2^{a_2} \cdots p_s^{a_s} \cdot L.$$

如果把 L 作素因子分解，可知上式右边 p_1 的方次至少为 a_1，而左边方次为 b_1. 由分解惟一性便知 b_1 至少为 a_1，即 $a_1 \leqslant b_1$. 同样有 $a_2 \leqslant b_2, \cdots, a_s \leqslant b_s$. 反过来，如果 $a_1 \leqslant b_1$，$a_2 \leqslant b_2, \cdots, a_s \leqslant b_s$，则 $N/M = p_1^{b_1 - a_1} p_2^{b_2 - a_2} \cdots p_s^{b_s - a_s}$ 是整数. 于是 M 整除 N，这样我们就证明了整除性判别法.

整除性判别法 如果正整数 M, N 有形如(*)式的素因子分解式，则 $M \mid N$ 的充分必要条件为 $a_1 \leqslant b_1, a_2 \leqslant b_2, \cdots, a_s \leqslant b_s$.

例如，我们求 198 的所有正整数因子，可以先求 198 的素因子分解式 $198 = 2 \cdot 3^2 \cdot 11$. 由上述判别法可知 198 的每个正整数因子均有 $2^a \cdot 3^b \cdot 11^c$ 的形式，其中 a 可取 0 或 1，b 可取 0，1 或 2，而 c 可取 0 或 1. 也就是说，a 和 c 有两种取法，而 b 有三种取法，所以 198 的正因子共有 $2 \cdot 2 \cdot 3 = 12$ 个，它们是

a	0	0	1	1	0	0	1	1	0	0	1	1
b	0	0	0	0	1	1	1	1	2	2	2	2
c	0	1	0	1	0	1	0	1	0	1	0	1
因子 $2^a 3^b 11^c$	1	11	2	22	3	33	6	66	9	99	18	198

现在仍设正整数 M 和 N 有分解式（*），我们求它们的最小公倍数和最大公因子．如上所述，M 的因子有形式 $L = p_1^{c_1} p_2^{c_2} \cdots p_s^{c_s}$，其中 $c_1 \leqslant a_1, \cdots, c_s \leqslant a_s$．如果 L 也是 N 的因子，则又有 $c_1 \leqslant b_1, \cdots, c_s \leqslant b_s$．所以 L 是 M 和 N 的公因子当且仅当对每个 $i = 1, 2, \cdots, s$，$c_i \leqslant a_i$ 且 $c_i \leqslant b_i$．而这条件又相当于 $c_i \leqslant \min\{a_i, b_i\}$（右边表示 a 和 b 当中最小者）．换句话说，$L = p_1^{c_1} \cdots p_s^{c_s}$ 是 M 和 N 的公因子当且仅当对每个 $i = 1, 2, \cdots, s$，$c_i \leqslant \min\{a_i, b_i\}$．如果求 M 和 N 的最大公因子 (M, N)，那么只需 $c_1, c_2, \cdots, c_s$ 取成满足上述条件的最大可能值即可，这个最大可能值显然是 $c_i = \min\{a_i, b_i\}$，这就表明

$$(M, N) = p_1^{c_1} p_2^{c_2} \cdots p_s^{c_s},$$

其中对每个 $i = 1, 2, \cdots, s$，$c_i = \min\{a_i, b_i\}$．例如，

$$120 = 2^3 \cdot 3 \cdot 5, \qquad 84 = 2^2 \cdot 3 \cdot 7$$

写成 $120 = 2^3 \cdot 3^1 \cdot 5^1 \cdot 7^0$，$84 = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^1$，可知

$$(120, 84) = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 = 12.$$

完全类似地推理可知，M 和 N 的最小公倍数为

$$[M, N] = p_1^{d_1} p_2^{d_2} \cdots p_s^{d_s},$$

其中 $d_i=\max\{a_i,b_i\}$（表示 a_i 和 b_i 的最大者）. 例如，$120=2^3\cdot3^1\cdot5^1\cdot7^0$，$84=2^2\cdot3^1\cdot5^0\cdot7^1$，最小公倍数为 $[120,84]=2^3\cdot3\cdot5\cdot7=840$.

利用整数的素因子分解式，我们还可以得出整数的许多性质. 下面几条性质是其中比较重要的，还有一些性质留给大家作为练习.

性质 1 设 a 和 b 为整数，p 为素数. 如果 $p\mid ab$，$p\nmid a$，则 $p\mid b$. 也就是说，如果乘积 ab 有素因子 p，而 a 没有素因子 p，则 b 必有素因子 p.

道理非常简单，如果 a 和 b 都没有素因子 p，那么 a 和 b 的素因子分解式中都不出现 p，于是乘起来便知 ab 的素因子分解式中也不出现 p，所以 p 不是 ab 的因子. 这就表明：若 p 是 ab 的因子，则 p 至少是 a 和 b 中某一个的因子. 也就是说，如果 p 不是 a 的因子，则 p 必是 b 的因子.

性质 2 若正整数 c 是正整数 a 和 b 的公因子$\left(\text{于是}\dfrac{a}{c},\dfrac{b}{c}\text{都是整数}\right)$，则 $(a,b)=c\cdot\left(\dfrac{a}{c},\dfrac{b}{c}\right)$，$[a,b]=c\cdot\left[\dfrac{a}{c},\dfrac{b}{c}\right]$.

证明 设 $a=p_1^{\alpha_1}\cdots p_s^{\alpha_s}$，$b=p_1^{\beta_1}\cdots p_s^{\beta_s}$，$c=p_1^{\gamma_1}\cdots p_s^{\gamma_s}$，则 $\dfrac{a}{c}=p_1^{\alpha_1-\gamma_1}\cdots p_s^{\alpha_s-\gamma_s}$，$\dfrac{b}{c}=$

$p_1^{\beta_1-\gamma_1}\cdots p_s^{\beta_s-\gamma_s}$，要证$(a,b)=c\left(\frac{a}{c},\frac{b}{c}\right)$，我们只需证明此式两边有相同的素因子分解式.对于每个素数$p_i(i=1,2,\cdots,s)$，左边分解式中p_i的指数为$\min(\alpha_i,\beta_i)$，而右边分解式中p_i的指数为$\gamma_i+\min(\alpha_i-\gamma_i,\beta_i-\gamma_i)$.所以，我们只需证明

$$\min(\alpha_i,\beta_i)=\gamma_i+\min(\alpha_i-\gamma_i,\beta_i-\gamma_i),$$

即要证$\min(\alpha_i,\beta_i)-\gamma_i=\min(\alpha_i-\gamma_i,\beta_i-\gamma_i)$.这个等式显然成立.不妨设$\alpha_i\leqslant\beta_i$，则$\alpha_i-\gamma_i\leqslant\beta_i-\gamma_i$，于是

$$\begin{aligned}\min(\alpha_i,\beta_i)-\gamma_i&=\alpha_i-\gamma_i\\&=\min(\alpha_i-\gamma_i,\beta_i-\gamma_i).\end{aligned}$$

这就证明了$(a,b)=c\cdot\left(\frac{a}{c},\frac{b}{c}\right)$.

同样可证$[a,b]=c\cdot\left[\frac{a}{c},\frac{b}{c}\right]$.

注 性质 2 表明$c\mid(a,b)$，也就是说：a和b的每个公因子c不仅小于最大公因子(a,b)，而且一定是最大公因子(a,b)的因子.而且性质 2 对于求a和b的最大公因子和最小公倍数是很有用的，因为如果我们知道了a和b的某个公因子c，我们可以把求(a,b)和$[a,b]$化为

求$\left(\frac{a}{c},\frac{b}{c}\right)$和$\left[\frac{a}{c},\frac{b}{c}\right]$,即化成求两个比较小的整数$\frac{a}{c}$,$\frac{b}{c}$的最大公因子和最小公倍数.比如说:

$$(120,84)=4\cdot(30,21)=4\cdot3\cdot(10,7)$$
$$=4\cdot3=12,$$
$$[120,84]=12\cdot\left[\frac{120}{12},\frac{84}{12}\right]=12\cdot[10,7]$$
$$=12\cdot70=840.$$

如果两个非零整数 a,b 的最大公因子为 1,即$(a,b)=1$(这也相当于说 a 和 b 没有公共素因子),我们称 a 和 b 是互素的.由于$(a,b)[a,b]=ab$,可知当 a 和 b 互素时,$[a,b]=ab$.

性质 3 设 a 和 b 是非零整数,$c=(a,b)$,则整数$\frac{a}{c}$和$\frac{b}{c}$互素.

直观上这个性质是容易理解的,因为$\frac{a}{c}$和$\frac{b}{c}$已经把 a 和 b 的最大公因子 c 抽走了,所以它们不可能再有大于 1 的公因子,即$\left(\frac{a}{c},\frac{b}{c}\right)=1$.

形式上的证明可以写成：$c=(a,b)=c\left(\frac{a}{c},\frac{b}{c}\right)$，两边除以 c，得到$\left(\frac{a}{c},\frac{b}{c}\right)=1$.

性质 4 设 a,b,c 均是非零整数. 如果 $c\mid ab$且$(c,a)=1$，则 $c\mid b$.

直观上这个性质也容易理解. 由于 c 和 a 互素，所以 c 的素因子均不在 a 的素因子分解式中. 但是$c\mid ab$，即 c 的素因子分解式必为 ab 的分解式的一部分，于是 c 的分解式必是 b 的一部分，即 $c\mid b$. 形式的证明可以写成：由于 c 是ab 的因子，又显然 c 是 bc 的因子，所以 c 是 ab 和 bc 的公因子. 根据性质 2 后面的注，可知 c 是(ab,bc)的因子. 但是$(ab,bc)=b(a,c)=b$(因为$(a,c)=1$)，因此 c 为 b 的因子，即$c\mid b$.

我们今后经常用到以上四条性质，所以特别地把证明写下来. 希望大家不要死记硬背这些性质，甚至它们的形式化的证明也不是最本质的，最本质的是这些性质的直观含义. 还有一些性质也是有用的，由于证明比较容易，留给大家作练习.

习　　题

1. 设 a,b,c 均是非零整数.

(1)如果 $a\mid b$，则 $a\mid bc$.

(2)如果 $a\mid b$，$b\mid c$，则 $a\mid c$.

(3)如果 $a\mid b$，$b\mid a$，则 $a=b$ 或者$a=-b$.

(4)如果 $a\mid b, a\mid c$,则 $a\mid(b\pm c)$.

2. 对每个整数 n,证明 $4\nmid(n^2+2)$.

3. 对每个奇数 n,证明 $8\mid(n^2-1)$.

4. 求证当 $n\geq 2$ 时,n^3+1 一定不是素数.

5. 求证对每个整数 a,$3\mid a^3-a$.

6. 求出满足$(a,b)=18,[a,b]=540$ 的整数 a 和 b.

7. 设 n,m 为正整数,求证:在 $n,2n,3n,\cdots,mn$ 当中共有(m,n)个是 m 的倍数.

作为算术基本定理和整除性质的一个应用,我们现在决定方程

$$x^2+y^2=z^2 \tag{2.2}$$

的所有正整数解.

设$(x,y,z)=(a,b,c)$是方程(2.2)的正整数解,则对任意正整数 n,将这组解同时乘以 n 之后,易知(na,nb,nc)也是一组正整数解.特别地,$(3n,4n,5n)(n=1,2,3,\cdots)$给出方程(2.2)的无穷多组正整数解.反过来,如果 d 是 a,b,c 的公因子,则易知$\left(\frac{a}{d},\frac{b}{d},\frac{c}{d}\right)$也是一组正整数解.如果 d 是 a,b,c 的最大公因子,则$\frac{a}{d}$,$\frac{b}{d}$ 和 $\frac{c}{d}$ 的公因子是 1.我们把最大公因子是 1 的正整数解称为方程(2.2)的基本解.由上所述,可知方程(2.2)的每个正整数解均可由某个基

本解同时乘上一个正整数而得到.所以我们只需找到方程(2.2)的全部基本解就可以了.下面定理给出方程(2.2)的全部基本解,定理表明基本解也有无穷多组.

定理 方程 $x^2+y^2=z^2$ 的全部基本解为

$$\begin{cases} x=m^2-n^2, \\ y=2mn, \\ z=m^2+n^2, \end{cases} \quad \text{或者} \quad \begin{cases} x=2mn, \\ y=m^2-n^2, \\ z=m^2+n^2, \end{cases} \tag{2.3}$$

其中 m 和 n 是两个互素的正整数,$m>n$,并且 $m+n$ 是奇数(即 m 和 n 一为奇数另一为偶数).

证 我们先证(2.3)式是 $x^2+y^2=z^2$ 的基本解.首先,$(m^2-n^2)^2+(2mn)^2=m^4-2m^2n^2+n^4+4m^2n^2=(m^2+n^2)$,并且再由 $m>n$ 可知(2.3)式是方程 $x^2+y^2=z^2$ 的正整数解.为证它是基本解,我们要证明每个素数 p 都不能是 m^2-n^2,$2mn$,m^2+n^2 的公因子.首先,由于 m 和 n 一奇一偶,可知 $m^2\pm n^2$ 为奇数,而 $2mn$ 为偶数,所以 2 不是它们的公因子.现在设奇素数 p 为 m^2-n^2,$2mn$ 和 m^2+n^2 的公因子.则 p 除尽$(m^2-n^2)+(m^2+n^2)=2m^2$ 和$(m^2+n^2)-(m^2-n^2)=2n^2$.由于

p 是奇素数,所以 p 除尽 m 和 n.这就与 m 和 n 互素假设相矛盾.这就表明(2.3)式给出的解是基本解.

现在我们证明方程 $x^2+y^2=z^2$ 的基本解一定有(2.3)式的形式,并且 m 和 n 一定满足定理中所述条件.设(x,y,z)是基本解,即 $x^2+y^2=z^2$,并且正整数 x,y,z 的最大公因子是1.事实上,x,y,z 当中任意两个整数的最大公因子都是 1,即它们是两两互素的.比如说,若 x 和 z 不互素,即某个素数 p 是 x 和 z 的公因子,则 $p\mid z^2-x^2=y^2$,于是 $p\mid y$.这表明 p 是 x,y,z 的公因子,与它们是基本解相矛盾.同样可知 x 和 y 互素,y 和 z 互素.由于 x,y,z 两两互素,可知 x 和 y 不能同时为偶数.另一方面,x 和 y 也不能同时是奇数.这是因为奇数的平方被 4 除余 1,所以若 x 和 y 都是奇数,则 $z^2=x^2+y^2$ 被 4 除余 2,但是 z^2 被 4 除不可能余 2,这就表明 x 和 y 一奇一偶.以下设 y 为偶数(从而 x 和 z 均为奇数),我们要证(x,y,z)必可表示成(2.3)式的第一种表达式(同样可证:若 x 为偶数,则(x,y,z)必可表示成(2.3)式的第二种表达式).

由于 y 为偶数,x 和 z 为奇数,因此$\frac{z\pm x}{2}$和$\frac{y}{2}$都是正整数(注意由 $x^2+y^2=z^2$ 可知 $z>$

x),并且

$$\frac{z+x}{2}\cdot\frac{z-x}{2}=\frac{z^2-x^2}{4}=\frac{y^2}{4}=\left(\frac{y}{2}\right)^2. \tag{2.4}$$

正整数$\frac{z+x}{2}$和$\frac{z-x}{2}$必然互素,因为若它们有公共素因子 p,则 p 也是$\frac{z+x}{2}+\frac{z-x}{2}=z$ 和 $\frac{z+x}{2}-\frac{z-x}{2}=x$的公因子,但我们已知 x 和 z 是互素的.这一矛盾表明$\frac{z+x}{2}$和$\frac{z-x}{2}$是互素的正整数.但(2.4)式表明它们的乘积是整数$\frac{y}{2}$的平方,利用算术基本定理可知$\frac{z+x}{2}$和$\frac{z-x}{2}$均是整数的平方,即

$$\frac{z+x}{2}=m^2,\quad \frac{z-x}{2}=n^2, \tag{2.5}$$

其中 m 和 n 是正整数,由$\frac{z+x}{2}$和$\frac{z-x}{2}$互素可知 m 和 n 互素.并且由(2.5)式可知 $z=m^2+n^2$, $x=m^2-n^2$,从而 $y^2=z^2-x^2=(m^2+n^2)^2-(m^2-n^2)^2=4m^2n^2$,即 $y=2mn$.最后由 $x>0$ 可知 $m>n$.这就完全证明了定理.

3 中国剩余定理

正如在上节定理的证明中看到的,我们常常要考虑整数相除的剩余.为此我们要引入一个方便的记号.设 m 是一个固定的正整数.我们称整数 a 和 b 模 m 是**同余**的,是指 $a-b$ 被 m 整除,即 $m \mid a-b$.这也相当于存在整数 c,使得 $a-b=cm$.我们把这件事表示成

$$a \equiv b(\bmod\ m),$$

其中 mod 是拉丁文 modulus(模)的字头.如果 $m \nmid a-b$,则称 a 和 b 模 m 不同余,写成 $a \not\equiv b(\bmod\ m)$.同余式写法与通常等式很相象.事实上,同余式具有等式的如下三条最基本性质.

性质 1(自反性) 每个整数均与自身模 m 同余,即 $a \equiv a(\bmod\ m)$.

性质 2(对称性) 若 a 与 b 模 m 同余,则 b 与 a 模的同余.即若 $a \equiv b(\bmod\ m)$,则 $b \equiv a(\bmod\ m)$.所以我们可以把它说成 a 和 b 模 m 彼此同余.

性质 3(传递性)　若 $a \equiv b \pmod{m}$，$b \equiv c \pmod{m}$，则 $a \equiv c \pmod{m}$.

这三条性质可由同余式的定义简单推出来(请读者自行练习).下面性质表明，同余式可以像等式一样地进行加减乘运算.

性质 4　若 $a \equiv b \pmod{m}$，$c \equiv d \pmod{m}$，则 $a \pm c \equiv b \pm d \pmod{m}$ 并且 $ac \equiv bd \pmod{m}$.

证明也不困难：若 $a \equiv b \pmod{m}$，$c \equiv d \pmod{m}$，则 $m \mid a-b$，$m \mid c-d$，所以

$$m \mid (a-b)+(c-d)=(a+c)-(b+d),$$

$$m \mid (a-b)-(c-d)=(a-c)-(b-d),$$

$$m \mid (a-b)c+b(c-d)=ac-bd.$$

这就表明 $a \pm c \equiv b \pm d \pmod{m}$ 和 $ac \equiv bd \pmod{m}$.

现在来谈除法.简单的例子会告诉你，同余式作除法要小心.例如 $2 \equiv 6 \pmod{4}$，即 $2.1 \equiv 2.3 \pmod{4}$，但是你不能把同余式两边用 2 去除，因为 $1 \equiv 3 \pmod{4}$ 是不对的.这是因为同余式 $ab \equiv ac \pmod{m}$ 相当于 $m \mid ab-ac=a(b-c)$，而同余式 $b \equiv c \pmod{m}$ 相当于 $m \mid b-c$.我们不能由 m 除尽 $a(b-c)$ 而推出 m 除尽其中一个因子 $b-c$.如果令 d 为 m 和 a 的最大公

因子(m,a),则由 $m\mid a(b-c)$可得到$\frac{m}{d}\Big|\frac{a}{d}(b-c)$.现在$\frac{m}{d}$和$\frac{a}{d}$是互素的整数,所以根据第2节的整除性质4,可知$\frac{m}{d}\Big|b-c$.即 $b\equiv c\left(\bmod \frac{m}{d}\right)$.也就是说,我们可以在某种程度上对同余式 $ab\equiv ac(\bmod m)$作除法.不过在消去 a 时,要把模 m 换成$\frac{m}{d}=\frac{m}{(m,a)}$才可.特别当$(a,m)=1$时,可以不改变模而得到 $b\equiv c(\bmod m)$.这是同余式除法与等式不同的一条重要性质,我们把它列成

性质5 若 $ab\equiv ac(\bmod m)$,则 $b\equiv c\left(\bmod \frac{m}{(m,a)}\right)$.特别当$(a,m)=1$时,$b\equiv c(\bmod m)$.

以上是同余式四则运算的最基本性质.简言之,你可以像通常所熟悉的等式那样进行加减乘运算,但是要记住同余式除法运算和等式有重大区别.由这些基本性质还可得到同余式其他一些运算性质.我们再列出重要的一条,其他留作习题.

性质6 设 m 是正整数,a 为整数.则存在整数 b 使得$ab\equiv 1(\bmod m)$的充分必要条件是$(a,m)=1$.

证 如果 $ab\equiv 1(\mathrm{mod}\ m)$，则 $ab-1=cm$，其中 c 为整数．现在令 $d=(a,m)$．则 d 除尽 a 和 m，从而也除尽 $ab-cm=1$，于是 $d=1$，即 $(a,m)=1$．反过来，假设 $(a,m)=1$．考虑 $m+1$ 个整数 $a^0(=1),a,a^2,\cdots,a^m$．我们知道，每个整数模 m 必同余于 $0,1,\cdots,m-1$ 这 m 个数当中的某一个，所以上述 $m+1$ 个整数当中一定有两个是模 m 同余的，即存在 a^i 和 a^j（其中 $i\neq j$），使得 $a^i\equiv a^j(\mathrm{mod}\ m)$．不妨设 $i>j$．则 $a^{i-j}\cdot a^j\equiv 1\cdot a^j(\mathrm{mod}\ m)$．由于已假定 $(a,m)=1$，可知 $(a^j,m)=1$．由性质 5 得到 $a^{i-j}\equiv 1(\mathrm{mod}\ m)$由 $i>j$ 知 $i-j-1$ 是非负整数．取 $b=a^{i-j-1}$，便得到 $ab=a^{i-j}\equiv 1(\mathrm{mod}\ m)$．这就证明了性质 6.

习　题

1．设 m 和 c 为正整数，a，b 为整数，则 $a\equiv b(\mathrm{mod}\ m)$当且仅当 $ac\equiv bc(\mathrm{mod}\ mc)$．

2．设 m 为正整数，a，b 为整数，$a\equiv b(\mathrm{mod}\,m)$则 $(a,m)=(b,m)$．

3．设 m 和 n 为正整数，a 和 b 为整数．则同余式 $a\equiv b(\mathrm{mod}\ n)$和 $a\equiv b(\mathrm{mod}\ m)$同时成立的充分必要条件是 $a\equiv b(\mathrm{mod}\ [m,n])$．

古代中国对整数同余性质有相当深入的研究．在公元 4 世纪左右写成的数学著作《孙子算

经》中记载了“物不知其数”问题.

今有物不知其数,三三数之剩二,五五数之剩三,七七数之剩二,问物几何?

这相当于求整数 x,使其满足以下的同余方程组

$$\begin{cases} x \equiv 2 \pmod 3, \\ x \equiv 3 \pmod 5, \\ x \equiv 2 \pmod 7). \end{cases}$$

程大位在《算法统宗》(1593 年)一书中把这个问题的解法总结成如下口诀:

三人同行七十稀,五树梅花廿一枝.

七子团圆半个月,除百零五便得知.

意思是说,将三个余数 2,3,2 分别乘上口诀前三句中提示的三个数 70,21 和 15,便得到一个解 $x=2\times 70+3\times 21+2\times 15=233$. 而最后一句的意思是说:将 233 减去 105 的任何倍数都是解. 口诀中的数 70,21,15 和 105 是怎么来的?用我们的同余式语言和性质,可以把这个问题的解法概括成如下的定理.

中国剩余定理 设 m_1, m_2, m_3 是两两互素的正整数,则对于任意三个整数 a_1, a_2, a_3,同余方程组

$$\begin{cases} x \equiv a_1 (\bmod\ m_1), \\ x \equiv a_2 (\bmod\ m_2), \\ x \equiv a_3 (\bmod\ m_3) \end{cases}$$

有解，并且若 $x = c$ 是一个解，则此同余方程组的全部整数解是模 $m_1 m_2 m_3$ 与 c 同余的全部整数.

证 由假设 m_1, m_2 和 m_3 两两互素，所以 $(m_2 m_3, m_1) = 1$，根据同余性质 6，存在整数 c_1，使得 $c_1 m_2 m_3 \equiv 1 (\bmod\ m_1)$，记 $d_1 = c_1 m_2 m_3$，则 d_1 满足

$$d_1 \equiv 1 (\bmod\ m_1),$$

$$d_1 \equiv 0 (\bmod\ n_2),$$

$$d_1 \equiv 0 (\bmod\ n_3).$$

换句话说，d_1 是 $m_2 m_3$ 的倍数，并且 d_1 模 m_1 同余于 1. 同样方法，可以找到 $m_1 m_3$ 的一个倍数 d_2，使得 $d_2 \equiv 1 (\bmod\ m_2)$，也可找到 $m_1 m_2$ 的一个倍数 d_3，使得 $d_3 \equiv 1 (\bmod\ m_3)$. 现在考虑整数 $c = a_1 d_1 + a_2 d_2 + a_3 d_3$，由于 d_1, d_2, d_3 的特性可知 c 模 m_1 同余于 a_1：

$$c \equiv a_1 \cdot 1 + a_2 \cdot 0 + a_3 \cdot 0 \equiv a_1 (\bmod\ m_1).$$

同样知 $c\equiv a_2 \pmod{m_2}$ 和 $c\equiv a_3 \pmod{m_3}$. 所以 $x=c$ 是同余方程组的一个解. 现在设 c' 是任意一个解, 则 $c'\equiv a_1 \pmod{m_1}$, $c\equiv a_1 \pmod{m_1}$, 于是 $c'-c\equiv 0 \pmod{m_1}$. 即 $c'-c$ 是 m_1 的倍数. 同样可知 $c'-c$ 也是 m_2 和 m_3 的倍数. 由于 m_1, m_2, m_3 两两互素, 这也相当于说 $c'-c$ 是 $m_1 m_2 m_3$ 的倍数, 即同余方程组的每个解均模 $m_1 m_2 m_3$ 同余于 c. 证毕.

在定理中取 m_1, m_2, m_3 分别为 3,5,7. 按照定理证明中给出的解法, 我们只需决定 d_1, d_2 和 d_3. d_1 是 $m_2 m_3 = 5\times 7 = 35$ 的倍数, 并且模 3 同余于 1, 不难看出 d_1 可取为 70. 类似方式可取 $d_2 = 21$ 满足 $d_2\equiv 0 \pmod{3\times 7}$ 和 $d_2\equiv 1 \pmod 5$, 可取 $d_3 = 15$ 满足 $d_3\equiv 0 \pmod{3\times 5}$ 和 $d_3\equiv 1 \pmod 7$. 于是得到"物不知其数"问题的一个解 $c = 2\times 70 + 3\times 21 + 2\times 15 = 233$. 最后 $3\times 5\times 7 = 105$, 根据定理可知模 105 同余于 233 的任何(正)整数都是"物不知其数"的解. 这就是程大位口诀中数 70,21,15 和 105 的秘密所在.

例 求 1999^{1999} 被 70 除所得的余数.

解 $70 = 2\times 5\times 7$. 令 $a = 1999^{1999}$. 我们容易求出 a 分别用 2,5,7 去除所得的余数:

$$a\equiv 1 \pmod 2,\quad a\equiv (-1)^{1999}\equiv -1 \pmod 5.$$

由于 $1999 \equiv 4 \pmod 7$，$4^3 \equiv 64 \equiv 1 \pmod 7$，所以

$$a \equiv 4^{1999} \equiv 4^{3 \cdot 666+1} \equiv 64^{666} \cdot 4 \equiv 4 \pmod 7.$$

于是 a 满足

$$\begin{cases} a \equiv 1 & \pmod 2, \\ a \equiv -1 & \pmod 5, \\ a \equiv 4 & \pmod 7. \end{cases}$$

对于中国剩余定理中 $m_1 = 2$，$m_2 = 5$，$m_3 = 7$ 的情形，算出 $d_1 = 35$，$d_2 = -14$，$d_3 = 50$. 因此

$$\begin{aligned} a &\equiv 1 \times 35 + (-1)(-14) + 4 \times 50 \\ &\equiv 249 \equiv 39 \pmod{70}, \end{aligned}$$

即 1999^{1999} 被 70 除的余数为 39.

本节的最后我们用整数的同余性质来求二元一次方程的整数解. 设 a 和 b 是非零整数，对每个整数 n，我们研究方程 $ax + by = n$ 的整数解 (x, y). 记 $d = (a, b)$. 如果此方程有整数解 (x, y)，则 $d \mid ax + by = n$，即 n 必需为 $d = (a, b)$ 的因子. 所以在 $(a, b) \nmid n$ 时，方程没有整数解. 现在设 $(a, b) \mid n$，则原方程化为

$$\frac{a}{(a, b)} x + \frac{b}{(a, b)} y = \frac{n}{(a, b)}.$$

这时$\frac{a}{(a,b)}$和$\frac{b}{(a,b)}$是互素的整数.下面定理表明在这种情形下方程一定有整数解,而且可写出它的全部整数解.

定理 3.1 设 a 和 b 是非零整数.

(1)对每个整数 n,方程

$$ax + by = n$$

有整数解的充分必要条件为$(a,b) \mid n$.

(2)如果$(a,b) = n$,并设 $x = c$ 和 $y = d$ 是此方程的一组整数解,则方程的全部整数解为

$$\begin{cases} x = c + \dfrac{b}{(a,b)}t, \\ y = d - \dfrac{a}{(a,b)}t, \end{cases}$$

其中 t 为任意整数.

证 我们已经证明了方程有整数解的必要条件为$(a,b) \mid n$.现在设$(a,b) = n$.令 $a' = \frac{a}{(a,b)}$, $b' = \frac{b}{(a,b)}$, $n' = \frac{n}{(a,b)}$,则原方程等价于 $a'x + b'y = n'$.由于$(a',b') = 1$,根据同余性质 6 可知同余方程 $a'x \equiv n' \pmod{b'}$有解.设整数 c 满足$a'c \equiv n' \pmod{b'}$,则 $a'c - n' = (-d) \cdot b'$,其中 d 为整数.于是$a'c + b'd = n'$,即$(x,y) = (c,d)$是方程的一组整数解.现在

对方程的任一整数解(x,y),则 $a'x+b'y=n'$.于是 $a'(x-c)+b'(y-d)=n'-n'=0$.于是 $a'(x-c)=-b'(y-d)$.从而 $b'\mid a'(x-c)$.但是$(a',b')=1$,因此$b'\mid x-c$即 $x-c=b't$(t 为整数).而 $a'b't=-b'(y-d)$,因此 $y-d=-a't$.于是必然

$$\begin{cases} x=c+b't, \\ y=d-a't. \end{cases}$$

易知对任意整数 t,这都是方程的整数解.这就完全证明了定理.

例 求方程 $123x+57y=531$ 的整数解.

解 由$(123,57)=3\mid 531$,可知方程必有整数解,用 3 去除方程两边得到 $41x+19y=177$.将此方程模 19 得到 $41x\equiv 177 \pmod{19}$,即 $3x\equiv 6 \pmod{19}$.于是 $x\equiv 2 \pmod{19}$,即 $x=2+19t$,而将 $x=2$ 代入 $41x+19y=177$,求出 $y=\frac{177-41\times 2}{19}=5$,所以$(x,y)=(2,5)$是方程组整解,而方程的全部整数解为

$$\begin{cases} x=2+19t, \\ y=5-41t, \end{cases}$$

其中 t 为任意整数.

4 同余类环和有限域

我们已经介绍了初等数论的最基本知识：整除性质和同余性质．这一小节是把这些知识作更深层次的思考和认识，引导和抽象出数学上的几个重要概念：群、环和域．事实上，这些概念的升华是在费马提出他的猜想之后完成的．费马猜想于1637年提出，而群的概念是法国数学家伽罗瓦（Galois，1811～1832）在19世纪初叶建立的．环和域的思想起源于18世纪欧拉和高斯等人对初等数论的研究，但是环和域理论和概念的深化也是19世纪之后的事情．我们以后将会看到，18世纪和19世纪人们对费马猜想的研究，对于环论和域论的发展起了重要的作用．而在本世纪费马猜想最终被证明，群论起了重大作用．

现在我们用已经讲过的初等数论知识来介绍群、环、域这三个概念．我们的叙述是从人们在日常生活中经常做的一件事情讲起，就是我

们常常要把各种事物进行某种方式的分类.

我们生活在一个五彩缤纷的世界之中,常常要把各种事物按照不同的标准和从不同的角度进行分类.事实上,合理的分类体现着人们对事物理解的加深和概念的升华.门捷列夫把近百种元素排来排去,最后发现了周期律.达尔文对生物的分类形成了生物进化的概念.古代人类把一只羊、一匹马和一块石头放在一类,并且把这一类起了个名字称为"1".两只羊不和一只羊放在一类,而和两块石头放在一类,起了个名字称为"2".数的概念便由此产生.

为了对分类进行严格的数学处理,需要建立适当的数学模型.一种常见的分类是由"等价关系"给出的.

假设我们要把某些事物进行分类,这些事物组成一个**集合** S,其中每个事物 a 称为集合 S 中的元素,表示成 $a\in S$.我们希望研究集合 S 的元素之间的一种关系,使得 S 中所有元素按这种关系分成一些类,同一类中的任意两个元素彼此都有这种关系,而不同类的任意两个元素都没有这种关系.这样的关系称为**等价关系**.比如说,地球上所有的人按着"性别相同"可以分成两类,这两类分别称为"男人"和"女人".所有男人之间性别相同,所有女人之间性别相同,而每个男人和每个女人性别不同.又比如

说，我们还可用“年龄相同”进行分类．同一类中任意两个人的年龄相同，而不同类的任意两个人的年龄不同．所以，上述两种分类都满足我们的条件，都是地球上所有人之间的等价关系．另一方面，如果把人们按照“年长”的关系对所有的人进行分类，我们无法使得每类当中任何两个人都有此关系．因为每个人和他自己属于同一类，但是他不比自己年长．又若 a 和 b 属于同一类，如果 a 比 b 年长，则 b 并不比 a 年长．所以“年长”关系不是等价关系．再比如“相互认识”也不是等价关系，因为若有三个人 a,b,c 属于一类，他们应当彼此都相互认识．但是若 a 和 b 认识，b 和 c 认识，而 a 和 c 可能相互不认识．

由上述可知，集合 S 上的一个关系 ~ 是等价关系，至少要满足以下三个条件（我们把 a 和 b 有关系表示成 $a \sim b$）．

(1)自反性：对于 S 中每个元素 a，$a \sim a$．

(2)对称性：设 $a,b \in S$．若 $a \sim b$，则 $b \sim a$．

(3)传递性：设 $a,b,c \in S$，若 $a \sim b$，$b \sim c$，则 $a \sim c$．

如果集合 S 中的一个关系满足这三项要求，我们就可把 S 中所有元素分类，使每个元素恰好属于一类．我们用$[a]$表示满足 $a \sim b$ 的

所有元素 b 组成的子集合.由于自反性 $a \sim a$,可知 a 属于子集合$[a]$.而对称性表明关系 $\sim$ 具有相互性:若 $b \in [a]$(即 $a \sim b$),则 $a \in [b]$(即 $b \sim a$).最后,由传递性可知每个类中任意两元素都彼此有关系 $\sim$,而不同类中的元素之间没有关系 $\sim$.所以,上述三条性质可作为等价关系的严格数学定义.

今后我们用 $\mathbb{Z}$ 表示所有整数表示的集合.我们考虑 $\mathbb{Z}$ 上的"模 m 同余"关系,其中 m 是固定的正整数.根据上节所述的前三条同余性质,可知这是等价关系.于是整数集合按此关系分成一些等价类.对每个整数 a,模 m 同余于 a 的所有整数组成一个等价类,记成$[a]$.于是$[a]=[b]$当且仅当 $a \equiv b \pmod{m}$.由于每个整数模 m 恰好同余于 $0,1,2,\cdots,m-1$ 当中的一个数,所以 $\mathbb{Z}$ 按模 m 同余共分成 m 个等价类$[0],[1],\cdots,[m-1]$.

学校里的学生分成一些班级,学校常常让每班派一位同学作代表在一起开会.同样地,在模 m 的 m 个同余类中每个同余类取一个整数,这样取出来的 m 个整数称为模 m 的**完全代表系**,简称为**完系**.所以,m 个整数 $a_1,\cdots,a_m$ 是模 m 的完系,是指这 m 个数模 m 彼此不同余,从而每个整数都恰好同余于它们当中的一个.例如,$0,1,2,\cdots,m-1$ 就是模 m 的一个完

系.又如,任意连续 m 个整数 $a, a+1, a+2, \cdots, a+m-1$ 都是模 m 的完系.再进一步有:

引理 4.1 (1)若 $\{a_1, \cdots, a_m\}$ 是模 m 完系,b 是与 m 互素的整数,则 $\{ba_1, \cdots, ba_m\}$ 也是模 m 完系.

(2)设 m 和 n 是互素的正整数,$\{a_1, \cdots, a_m\}$ 和 $\{b_1, \cdots, b_n\}$ 分别是模 m 和模 n 的完系,则 $\{na_i + mb_j: (1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n)\}$ 是模 mn 的完系.

证 (1)我们只需证明 $ba_1, \cdots, ba_m$ 这 m 个数彼此模 m 不同余.如果 $ba_i \equiv ba_j \pmod{m}$,由 b 与 m 互素知 $a_i \equiv a_j \pmod{m}$.再由 $\{a_1, \cdots, a_m\}$ 为完系可知 $i=j$,从而当 $i \neq j$ 时,$ba_i \not\equiv ba_j \pmod{m}$,即 $ba_1, \cdots, ba_m$ 为模 m 完系.

(2)我们只需证 mn 个数 $na_i + mb_j (1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n)$ 彼此模 mn 不同余.如果 $na_i + mb_j \equiv na_{i'} + mb_{j'} \pmod{mn}$,则这两个数模 m 也同余,于是 $na_i \equiv na_{i'} \pmod{m}$.由假设 n 和 m 互素,可知 $a_i \equiv a_{i'} \pmod{m}$.再由 $\{a_1, \cdots, a_m\}$ 是模 m 的完系,可知 $i = i'$.同样可证 $j = j'$.所以当 $i \neq i'$ 或者 $j \neq j'$ 时,$na_i + mb_j$ 和 $na_{i'} + mb_{j'}$ 模 mn 不同余.即 $na_i + mb_j (1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n)$ 是模 mn 的完系.

现在做一件看上去简单但是很重要的事情.我们已经把全体整数 $\mathbb{Z}$ 按模 m 分成 m 个同余类$[0],[1],\cdots,[m-1]$.现在把每个同余类$[a]$不再看成是 $\mathbb{Z}$ 的一个子集合,而看成一个抽象的**元素**.而 m 个元素 $[0],[1],\cdots,[m-1]$组成一个新的集合,表示成 $\mathbb{Z}_m$.例如对 $m=2$,我们不把$[0]$看成是所有被 2 除尽的整数组成的集合,而看成一个元素:"偶数",而$[1]$看成一个元素,叫"奇数".于是 $\mathbb{Z}_2=\{[0],[1]\}$是由偶数和奇数这两个元素构成的集合.一般地,若 $a_1,\cdots,a_m$ 是模 m 的任何一个完系,则 $\mathbb{Z}_m=\{[a_1],\cdots,[a_m]\}$.例如 $\mathbb{Z}_5=\{[-3],[0],[1],[8],[4]\}$.由于 $a\equiv b \pmod m$相当于$[a]=[b]$,所以整数之间的模 m 同余式相当于集合 $\mathbb{Z}_m$ 中的等式.

以上我们把整数集合 $\mathbb{Z}$ 分成模 m 的 m 个同余类.将每个同余类看作一个元素,得到 m 个元素的同余类集合 $\mathbb{Z}_m$.在整数集合 $\mathbb{Z}$ 中是有加、减乘法运算的.我们是否可以把 $\mathbb{Z}$ 中这些运算自然地引到 $\mathbb{Z}_m$ 中来呢? 所谓自然的方式就是对 a 和 b 所属的同余类$[a]$和$[b]$,定义它们的和是 $a+b$ 所属的同余类,即$[a]+[b]=[a+b]$.但是要注意:同余类可以取不同的代表元素.假如同余类$[a]$和$[b]$中又分别取了另外

代表元素 c 和 d，即 $[a]=[c]$，$[b]=[d]$。如果 $[a+b]$ 不等于 $[c+d]$，如此定义的同余类上加法运算就没有道理了。幸亏同余式性质保证 $[a+b]=[c+d]$，因为由 $[a]=[c]$ 和 $[b]=[d]$ 可知 $a\equiv c \pmod{m}$，$b\equiv d \pmod{m}$。于是 $a+b\equiv c+d \pmod{m}$，这表明 $[a+b]=[c+d]$。同样地，我们可以定义

$$[a]-[b]=[a-b],\quad [a]\cdot[b]=[ab].$$

数学上把具有加、减、乘三种运算，并且满足通常的结合律、交换律和分配律的集合，称为一个**环**。比如整数集合 $\mathbb{Z}$ 是环，叫**整数环**。现在 $\mathbb{Z}_m$ 也是环，叫**同余类环**。在环 $\mathbb{Z}_m$ 中元素 $[0]$ 起着"零"的作用，即 $[0]+[a]=[a]$。而元素 $[1]$ 起着整数 1 的作用，即 $[1]\cdot[a]=[a]$。此外有

$$2[a]=[a]+[a]=[a+a]=[2a],$$

$$-[a]=[-a].$$

由此可知对每个整数 n，均有 $n[a]=[na]$。注意在 $\mathbb{Z}_m$ 中，$m[a]=[ma]=[0]$，即每个元素的 m 倍均为零，这是 $\mathbb{Z}_m$ 与整数环 $\mathbb{Z}$ 的重要区别。

环 $\mathbb{Z}_m$ 和 $\mathbb{Z}$ 的最重要区别是在做除法上。让我们重新考查一下上节的同余性质 6。这个性

质是说：对于整数 a，存在整数 b 使得 $ab \equiv 1 \pmod{m}$ 当且仅当 a 和 m 互素．用环 $\mathbb{Z}_m$ 中的语言，$ab \equiv 1 \pmod{m}$ 就是 $[a]\cdot[b]=[1]$．即 $[a]$ 是环 $\mathbb{Z}_m$ 中的可逆元素，$[a]$ 的逆就是 $[b]$，表示成 $[a]^{-1}$．再考查上节的习题 2，这个习题是说：如果 $a \equiv b \pmod{m}$，则 $(a,m)=(b,m)$．所以若 a 与 m 互素，则 $[a]$ 中每个元素 b 均与 m 互素，我们可以说同余类 $[a]$ 与 m 互素．于是，同余性质 6 用环 $\mathbb{Z}_m$ 中的语言可以说成：$[a]$ 是可逆元素当且仅当 $[a]$ 与 m 互素．

我们可以用可逆元素去除任何元素，因为若 $[a]$ 可逆，则 $[c]/[a]$ 就是 $[c]\cdot[a]^{-1}$．比如对 $m=10$，一共有 10 个模 10 同余类，其中与 10 互素的有 4 个：$[1]$，$[3]$，$[7]$ 和 $[9]$．它们是 $\mathbb{Z}_{10}$ 中可逆元素．

$$[1]^{-1}=[1],\quad [3]^{-1}=[7],$$

$$[7]^{-1}=[3],\quad [9]^{-1}=[9].$$

这是因为 $1\cdot1 \equiv 3\cdot7 \equiv 9\cdot9 \equiv 1 \pmod{10}$．于是可用 $[3]$ 去除任何同余类 $[a]$．比如在 $\mathbb{Z}_{10}$ 中

$$[4]/[3]=[4]\cdot[3]^{-1}=[4][7]=[28]=[8].$$

如果你对这种运算仍不习惯，可以回到同余式运算上来：

$$4/3 \equiv \frac{4+20}{3} \equiv \frac{24}{3} \equiv 8 \pmod{10},$$

即$[4]/[3]=[8]$.所以同余式也可写成分数形式,并且分子分母均可随意加上 m 的倍数,逐渐把分母消去.只是要保证分母一定要与 m 互素.比如在 $\mathbb{Z}_{28}$ 中,$[15]$与 28 互素.我们求$[22]/[15]$:

$$\begin{aligned}\frac{22}{15} &\equiv \frac{22+28}{15} \equiv \frac{50}{15}\\ &\equiv \frac{10}{3} \equiv \frac{10-28}{3}\\ &\equiv -\frac{18}{3} \equiv -6\\ &\equiv 22 \pmod{28}.\end{aligned}$$

于是$[22]/[15]=[-6]=[22]$.

一个环中可逆元素愈多,在这个环中做除法就愈灵活.在整数环 $\mathbb{Z}$ 中只有两个可逆元素:± 1.而在同余类环 $\mathbb{Z}_m$ 中可以有较多的可逆元素.比如 $\mathbb{Z}_{10}$ 的 10 个元素当中就有 4 个可逆元素.特别当 m 是素数 p 时,$\mathbb{Z}_p$ 中 p 个同余类$[0],[1],\cdots,[p-1]$当中,后 $p-1$ 个同余类均与 p 互素,即除了$[0]$之外,其余均是可逆元素.换句话说,除了零元素不能做为除数之外,

$_p$ 中任何两个元素都可做除法.这样的环就称为是一个**域**.因此对每个素数 p，$_p$ 是 p 个元素的域.整数环 不是域，但是我们知道，所有有理数全体可做加减乘运算，并且除了 0 不能做除数之外，任何两个有理数相除仍是有理数，所以全体有理数组成一个域，称为有理数域，表示成 .同样的，所有实数是一个域，所有复数也是域，分别称为实数域和复数域，表示成 和 .这些域上的四则运算我们在中学已经熟悉.现在对每个素数 p，我们构造出有限个(p 个)元素的域 $_p$，称为**有限域**，有限域 $_p$ 通常也表示成 $_p$.今后我们讲到费马猜想的进展时，要考虑这些域以及其他的域(代数数域).

现在我们举两个例子，这些例子虽然内容都是关于同余式的论断，但是我们希望大家能够习惯于使用同余类环和有限域的语言.比如说:若在某个环中有等式 $ab = ac$，并且 a 是此环中的可逆元素，则可用 a 去除两边得到 $b = c$.

例 1 若 p 是素数，a 是与 p 互素的整数，则 $a^{p-1} \equiv 1 \pmod p$.这个结果是费马的一个猜想，由欧拉给出证明并且加以推广(见后面例 3)，称为费马小定理.

证 $\{0,1,2,\cdots,p-1\}$ 是模 p 的一个完

系，由于 a 与 p 互素，根据引理 4.1 可知 $\{0, a, 2a, \cdots, (p-1)a\}$ 也是模 p 的完系．也就是说，$[0],[1],\cdots,[p-1]$ 是有限域 $\mathbb{F}_p$ 中全部 p 个元素，而 $[0],[a],[2a],\cdots,[(p-1)a]$ 也是 $\mathbb{F}_p$ 的全部 p 个元素，只是排列次序可能不同．除掉 $[0]$ 之外，我们就有

$$\begin{aligned}[1][2]\cdots[p-1] &= [a][2a]\cdots[(p-1)a] \\ &= [1][a][2][a]\cdots[p-1][a] \\ &= [1][2]\cdots[p-1][a]^{p-1} \\ &= [1][2]\cdots[p-1][a^{p-1}].\end{aligned}$$

由于非零元素 $[1],[2],\cdots,[p-1]$ 都是域 $\mathbb{F}_p$ 中可逆元素，所以上面等式两边可消去这些元素，得到 $[1]=[a^{p-1}]$，这就相当于 $a^{p-1}\equiv 1 \pmod p$．

例 2 设 p 是奇素数．则存在整数 a 使得 $a^2\equiv -1 \pmod p$ 当且仅当 $p\equiv 1 \pmod 4$．

证 每个奇素数 p 模 4 同余于 1 或 3．先设 $p\equiv 1 (\mathrm{mod}\ 4)$．我们要找到有限域 $\mathbb{F}_p$ 中一个元素 $[a]$，使得 $[a]^2=[-1]=-[1]$．为此我们考虑 $\mathbb{F}_p$ 中所有 $p-1$ 个非零元素的乘积

$$A=[1][2]\cdots[p-1].$$

一方面，由于 $[p-1]=[-1]=-[1]$，$[p-2]$

$= -2, \cdots, \left[\frac{p+1}{2}\right] = -\left[\frac{p-1}{2}\right]$. 可知 $\left(注意\frac{p-1}{2}为偶数\right)$

$$A = [1][2]\cdots\left[\frac{p-1}{2}\right]\cdot\left(-\left[\frac{p-1}{2}\right]\right)\cdots(-[2])(-[1])$$

$$= (-1)^{\frac{p-1}{2}}[1]^2[2]^2\cdots\left[\frac{p-1}{2}\right]^2 = [a]^2,$$

其中 $a = 1\cdot 2\cdots\left(\frac{p-1}{2}\right)$. 另一方面,对每个 $1 \leqslant a \leqslant p-1$, $[a]$ 均是 $\mathbb{Z}_p$ 中可逆元素,设 $[a]^{-1} = [b]$,则 $[a]$, $[b]$ 彼此互逆. 但是可能 $[a] = [b]$,即 $[a]$ 自身为 $[a]$ 的逆, $[a] = [a]^{-1}$,也就是 $[a][a] = 1$,即 $[a^2] = [1]$. 这时必然 $[a^2 - 1] = [0]$,即 $[a+1][a-1] = [0]$. 注意在域中若 $xy = 0$ 则必然 x, y 当中有一个为 0(因若 $x \neq 0$,则 x 可逆,从而 $xy = 0 = x\cdot 0$ 可消去 x 得到 $y = 0$). 于是由 $[a+1][a-1] = [0]$ 得到 $[a+1] = [0]$ 或者 $[a-1] = [0]$,即 $[a] = [1]$ 或者 $[a] = [-1]$. 所以满足 $[a] = [a]^{-1}$ 的只有 $[a] = [1]$ 和 $[a] = [-1] = [p-1]$ 这两个元素. 而剩下 $p-3$ 个非零元素一定分成 $\frac{p-3}{2}$ 对,每对的两个元素彼此互逆,从而乘起来为 $[1]$. 这就表明 $A = [1][2]\cdots[p-2][p-1] = [1]\cdot$

$[p-1]=[-1]$.但是$[A]=[a]^2$,于是$[a]^2=[-1]$.从而当 $p\equiv 1(\mathrm{mod}\ 4)$时我们有整数 a 使得 $a^2\equiv -1(\mathrm{mod}\ p)$.

现在对 $p\equiv 3(\mathrm{mod}\ 4)$情形,要证不存在整数 a 使得 $a^2\equiv -1(\mathrm{mod}\ p)$.我们用反证法,假设在有限域 $\mathbb{Z}_p$ 中$[a]^2=[-1]$,则$[a]^4=[-1]^2=[1]$.进而由费马小定理(例 1)知道$[a]^{p-1}=[1]$.但是 $p-1\equiv 3-1\equiv 2(\mathrm{mod}\ 4)$,即 $p-1$ 被 4 除余 2,所以 $p-1=4l+2$,其中 l 为正整数.于是

$$
\begin{aligned}
[1] &= [a]^{p-1} = [a]^{4l+2} \\
&= ([a]^4)^l \cdot [a]^2 \\
&= [a]^2 = [-1].
\end{aligned}
$$

但是$[1]=[-1]$在 $\mathbb{Z}_p$ 中不可能成立,因为对于奇素数 p,$1\not\equiv -1(\mathrm{mod}\ p)$.这个矛盾表明当 $p\equiv 3(\mathrm{mod}\ 4)$时不存在整数 a 使得 $a^2\equiv -1(\mathrm{mod}\ p)$.

如上所述,环 $\mathbb{Z}_m$ 中有许多可逆元素.我们以 $\varphi(m)$表示 $\mathbb{Z}_m$ 中可逆元素的个数,$\varphi(m)$称为**欧拉函数**.现在我们要给出计算 $\varphi(m)$的一个公式.为了做这件事,我们可取模 m 的一个完系$\{a_1,a_2,\cdots,a_m\}$,其中每个与 m 互素的 a_i,$[a_i]$是 $\mathbb{Z}_m$ 中可逆元素,所以 $a_1,a_2,\cdots,a_m$ 中

与 m 互素的 a_i 共 $\varphi(m)$ 个. 这 $\varphi(m)$ 个 a_i 称为组成模 m 的一个**缩系**. 特别取模 m 的完系 $\{1,2,3,\cdots,m\}$, 则 $\varphi(m)$ 就是 $1,2,\cdots,m$ 当中与 m 互素的数的个数. 而一般情况下, $\varphi(m)$ 个整数 $c_1,c_2,\cdots,c_{\varphi(m)}$ 是模 m 的缩系, 即是指它们均与 m 互素并且彼此模 m 不同余. 因为这时 $[c_1],[c_2],\cdots,[c_{\varphi(m)}]$ 就是环 $\mathbb{Z}_m$ 的全部 $\varphi(m)$ 个不同的可逆元素. 为了给出计算 $\varphi(m)$ 的公式, 我们有与引理 4.1 类似的如下结果.

引理 4.2 (1) 若 $\{a_1,\cdots,a_s\}$ 是模 m 的缩系 $(s=\varphi(m))$, 则对每个与 m 互素的整数 b, $\{ba_1,\cdots,ba_s\}$ 也是模 m 的缩系.

(2) 设 m 和 n 是互素的正整数, $\{a_1,\cdots,a_s\}$ 和 $\{b_1,\cdots,b_t\}$ 分别是模 m 和模 n 的缩系 (于是 $s=\varphi(m)$, $t=\varphi(n)$), 则 $\{na_i+mb_j(1\leqslant i\leqslant s),(1\leqslant j\leqslant t)\}$ 是模 mn 的缩系. 于是 $\varphi(mn)=\varphi(m)\varphi(n)$.

证 (1) 与引理 4.1 一样可证 $ba_1,\cdots,ba_s$ 模 m 彼此互不同余. 由于 b 和 a_i 均与 m 互素, 所以 ba_i 也与 m 互素. 所以 $s=\varphi(m)$ 个整数 $\{ba_1,\cdots,ba_s\}$ 是模 m 的缩系.

(2) 与引理 4.1 一样可证 st 个数 na_i+mb_j $(1\leqslant i\leqslant s,1\leqslant j\leqslant t)$ 模 m 彼此互不同余. 进而, 由于 $(a_i,m)=(n,m)=1$ 可知

$$(na_i + mb_j, m) = (na_i, m) = 1.$$

同样有$(na_i + mb_j, n) = (mb_j, n) = 1$.于是$(na_i + mb_j, mn) = 1$.所以 st 个数 $na_i + mb_j$ 均与 mn 互素.为了证明它们是模 mn 的缩系,我们还需证明这 st 个数给出与 m 互素同余类的全部代表元,即要证对与 mn 互素的每个整数 x,均有 i 和 $j(1 \leqslant i \leqslant s, 1 \leqslant j \leqslant t)$,使得 $x \equiv na_i + mb_j \pmod{mn}$.证明如下:由于 n 和 m 互素,所以由已证明的(1)可知$\{na_1, \cdots, na_s\}$是模 m 的缩系.因此有 i 使得 $x \equiv na_i \pmod{m}$.同样可知有 j 使得 $x \equiv mb_j \pmod{n}$.于是

$$x \equiv na_i + mb_j \pmod{m},$$

$$x \equiv na_i + mb_j \pmod{n}.$$

由$(m, n) = 1$ 可知 $x \equiv na_i + mb_j \pmod{mn}$.这就最终证明了 st 个数 $na_i + mb_j$ 构成模 mn 的缩系.于是 $\varphi(mn) = st = \varphi(m)\varphi(n)$.

现在可以给出 $\varphi(m)$的计算公式.根据引理4.2,若 m, n 是互素的正整数,则 $\varphi(m, n) = \varphi(m)\varphi(n)$.特别有 $\varphi(m) = \varphi(m)\varphi(1)$,所以 $\varphi(1) = 1$.若 $m \geqslant 2$,令

$$m = p_1^{a_1} \cdots p_s^{a_s},$$

其中 $p_1, \cdots, p_s$ 是不同的素数,$a_1, \cdots, a_s$ 均为

正整数，则 $p_1^{a_1},\cdots,p_s^{a_s}$ 彼此互素，所以由引理 4.2 可知

$$\varphi(m)=\varphi(p_1^{a_1})\cdots\varphi(p_s^{a_s}).$$

这就把问题归结为计算 $\varphi(p^a)$，其中 p 为素数而 $a\geqslant 1$. 但是 $\varphi(p^a)$ 是 $1,2,\cdots,p^a$ 当中与 p^a 互素的数的个数. 也就是 p^s 个数 $1,2,\cdots,p^s$ 当中不被 p 除尽的数有多少个. 由于被 p 除尽的有 $p^s/p=p^{s-1}$ 个. 所以不被 p 除尽的有 $p^s-p^{s-1}=p^{s-1}(p-1)$ 个. 即 $\varphi(p^s)=p^{s-1}(p-1)$. 于是我们最后得到

引理 4.3(欧拉) $\varphi(1)=1$. 并且对每个整数 $m\geqslant 2$，设 $m=p_1^{a_1}\cdots p_s^{a_s}$，其中 $p_1,\cdots,p_s$ 是不同的素数，$a_1,\cdots,a_s$ 均为正整数，则

$$\begin{aligned}\varphi(m)&=p_1^{a_1-1}\cdots p_s^{a_s-1}(p_1-1)\cdots(p_s-1)\\&=m\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_s}\right).\end{aligned}$$

例如对 $m=45=3^2\cdot 5$，则 $\varphi(45)=45\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)=24$，即环 $\mathbb{Z}_{45}$ 中共有 24 个可逆元素.

下面的结果是欧拉对费马小定理所作的推广.

例 3(欧拉定理) 若 a 是与 m 互素的整

数,则

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

证 和证费马小定理(例 1)相仿,设 $s=\varphi(m)$,$\{a_1,\cdots,a_s\}$为模 m 的缩系,由于 a 和 m 互素,可知$\{aa_1,\cdots,aa_s\}$也是模 m 的缩系(引理 4.2).于是$[a_1],\cdots,[a_s]$为环 $\mathbb{Z}_m$ 中 $s=\varphi(m)$个可逆元素,$[aa_1],\cdots,[aa_s]$也是$\mathbb{Z}_m$中这 $\varphi(m)$个可逆元素.所以

$$\begin{aligned}[a_1]\cdots[a_s] &= [aa_1]\cdots[aa_s]\\ &= [a^s][a_1]\cdots[a_s].\end{aligned}$$

由于$[a_1],\cdots,[a_s]$均可逆,因此在上式中可把它们消去,得到$[1]=[a^s]$$(s=\varphi(m))$,即$a^{\varphi(m)}\equiv 1 \pmod{m}$.

注记 1 用欧拉定理可知:若 a 与 m 互素,则在环 $\mathbb{Z}_m$ 中$[a]$为可逆元素,并且由$[a]^{\varphi(m)}=1$ 可知$[a]$的逆为$[a]^{\varphi(m)-1}$.例如对 $m=45$,$[14]$是$\mathbb{Z}_{45}$中可逆元素.由 $\varphi(45)=24$ 可知$[14]^{-1}=[14^{23}]$.但通常当 m 很大时,$a^{\varphi(m)-1}$很大,不如下面方法方便:

$$\frac{1}{14} \equiv \frac{46}{14} \equiv \frac{23}{7} \equiv 3+\frac{2}{7}$$

$$\equiv 3+\frac{2-3\cdot 45}{7} \pmod{45}$$

$$\equiv 3 + \frac{2 - 3 \cdot 3}{7} - 18$$

$$\equiv -16 \equiv 29 \pmod{45}$$

从而在 $\mathbb{Z}_{45}$ 中 $[14]^{-1} = [29]$，即 $14 \cdot 29 \equiv 1 \pmod{45}$.

以上我们通过具体例子介绍了环和域的概念（整数环 $\mathbb{Z}$，同余类环 $\mathbb{Z}_m$，有理数域 $\mathbb{Q}$，实数域 $\mathbb{R}$，复数域 $\mathbb{C}$ 和有限域 $\mathbb{Z}_p = \mathbb{F}_p$）. 这些代数结构中本质上有两种运算：加法和乘法. 而减法是加法的逆运算，除法是乘法的逆运算，并且除法只对一部分情形（例如用可逆元素去除）才可以进行. 加法和乘法不是相互孤立的，要满足分配律. 在本节的最后我们再介绍第三种代数结构：群，这种结构比环和域要简单，只需要一种代数运算.

定义 集合 G 对于运算·称为是群，是指：

(1)（结合律） 对 G 中任意三个元素 a, b, c,

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

(2) 集合 G 中存在（惟一）元素 e，使得对 G 中每个元素 a，均有 $a \cdot e = e \cdot a = a$. 称 e 为**幺元素**.

(3)集合 G 中每个元素 a 均有（惟一）元素 b，使

得 $a \cdot b = b \cdot a = e$. 称 b 为 a 的**逆元素**，表示成 a^{-1}.

注意我们并没有假定运算满足交换律. 如果运算满足交换律，即对 G 中任意元素 a 和 b，均有 $a \cdot b = b \cdot a$，则 G 称为**交换群**.

我们在前面已经见到了许多群的例子.

(1) 整数全体 $\mathbb{Z}$ 对于加法形成群，幺元素即是 $0(0 + a = a)$，而整数 a 的逆元素为 $-a(a + (-a) = 0)$. 同余类环 $\mathbb{Z}_m$ 对于加法也是群，幺元素为$[0]$，$[a]$的逆元素为$[-a]$. 一般地，每个环对于加法均是群.

(2) 由于环对于乘法不一定有逆运算，所以环对乘法不一定是群. 但是我们若只考虑环中的可逆元素，以 $R^\times$ 表示环 R 中可逆元素全体组成的集合，由于可逆元素相乘仍可逆，可逆元素 a 的逆元素 a^{-1} 也可逆(a^{-1}的逆元素就是 a). 可知 $R^\times$ 对于乘法为群. 例如整数环 $\mathbb{Z}$ 的可逆元素只有 ± 1，所以$\{1, -1\}$对乘法为群. 模 m 同余类环 $\mathbb{Z}_m$ 的 $\varphi(m)$个可逆元素形成乘法群$(\mathbb{Z}_m)^\times$，幺元素为$[1]$. 对于每个域 F，由于 F 中非零元素均是可逆元素，所以 $F^\times$ 就是 F 去掉零元素 0，即每个域 F 的非零元素全体 $F^\times$ 对乘法是群. 例如非零有理数乘法群 $\mathbb{Q}^\times$，非零实数乘法群 $\mathbb{R}^\times$ 等等. 有限域 $\mathbb{F}_p$ 的 $p-1$ 个非

零元素[1],[2],…,[$p-1$]也形成乘法群$\mathbb{F}_p^*$.

以上例子中的群都是交换群.最后我们举一个非交换群的例子,这个群在今后讲模形式理论时要用到.

我们把任意4个整数a,b,c,d排成如下的方阵

$$M=\begin{pmatrix}a & b\\ c & d\end{pmatrix}.$$

这个方阵的行列式定义成

$$|M|=\begin{vmatrix}a & b\\ c & d\end{vmatrix}=ad-bc.$$

方阵M和方阵

$$N=\begin{pmatrix}a' & b'\\ c' & d'\end{pmatrix}$$

的乘法定义为

$$MN=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}a' & b'\\ c' & d'\end{pmatrix}=\begin{pmatrix}aa'+bc' & ab'+bd'\\ ca'+dc' & cb'+dd'\end{pmatrix}.$$

可以直接验证,MN的行列式为M和N的行列式之乘积,即$|MN|=|M|\cdot|N|$.特别地,如果M和N的行列式均为1,则MN的行列式

也为 1.所以我们令 G 表示行列式为 1 的所有方阵

$$M=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\quad(a,b,c,d\in\mathbb{Z},ad-bc=1)$$

构成的集合,那么 G 中可以进行上述乘法运算.可以验证这个乘法运算满足结合律,并且具有幺元素 $I=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$,因为

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}.$$

进而,对于集合 G 中每个方阵 $M=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$,则方阵 $N=\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$ 也属于 G,并且易知 $MN=NM=I$.所以 N 就是 M 的逆方阵.综合上述,可知 G 对上述乘法是群.这个群不是交换群.因为对于 G 中方阵 $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ 和 $\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$,我们有

$$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}=\begin{pmatrix}-1 & 1\\ -1 & 0\end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

以上我们粗浅地介绍了群、环、域的概念和一些具体例子.在讲到费马猜想发展的时候会用到这些概念和介绍更多的例子.

5 费马猜想

古希腊数学从公元 3 世纪之后开始衰落,一直到公元 14 世纪,欧洲处于教会的黑暗统治之下,数学和整个科学进展缓慢.在 15 和 16 世纪欧洲文艺复兴时代,数学也得到复兴和发展,但主要是基于航海、天文、建筑和绘画等需要的画法几何学,数论的进展不大.17 世纪和 18 世纪数论才取得很大的发展,当时的数论中心在法国,大数论学家有法国人拉普拉斯、勒让德、拉格朗日和费马等,惟一的例外欧拉不是法国人.

古希腊数学家丢番图于公元 3 世纪所写的数论名著《算术》在 1621 年被翻译成拉丁文.1637 年,费马在阅读此书中讨论方程 $x^2+y^2=z^2$ 的那一页的空白处写了如下的评注:

但是一个立方数不能分拆成两个立方数,一个四次方数不能分拆成两个四次方数.一般说来,除平方之外,任何次幂不能分拆成两个同

次幂.我发现了一个真正奇妙的证明,但书上的空白太小,写不下.

用确切的数学语言来说,费马声称他证明了:对于每个大于2的正整数 n,方程 $x^n+y^n=z^n$ 都没有正整数解.

事实上,人们只看到费马对 $n=4$ 的情形给出的证明.经过350年的努力,怀尔斯于1994年才第一个证明了费马的这个论断.所以我们宁愿将它称为费马猜想.

皮埃尔·德·费马(Pierre de Fermat)在1601年8月20日出生于法国西南部一个小镇.父亲是富有的皮革商人.1631年费马获奥尔良大学民法学士学位后从事行政工作,后以律师为职业.但从事数学研究是他最大的爱好.他是解析几何、微积分和概率论的先驱者之一,但他对数学发展的最大贡献是在数论方面.他的数论研究全都写在给朋友的通信中和对丢番图《算术》一书的批注中.他叙述了许多数论结果,但绝大部分均未给出证明.他在《算术》一书中除了上面的著名猜想之外,还在空白处写到:

1. 形如 $4n+3$ 的素数不能表成二平方和.

2. 形如 $4n+1$ 的素数能表成二平方和,并且本质上只有一种方式.

1640年,他在给朋友的信中说:

3. 若 n 不是素数,则 2^n-1 也不是素数.



图 1

4. 若 p 为素数，则对每个整数 a，a^p-a 可被 p 整除（现在这称为费马小定理）.

5. 若 p 为奇素数，则 2^p-1 的素因子均有形式 $2kp+1$.

6. 形如 $2^{2^n}+1$ 的整数（$n=0,1,2,\cdots$）都是素数.

……

费马在 1665 年去世，他的这些信件和评注由他的长子于 1670 年收集出版. 80 年之后，欧拉研究费马的工作，引起对数论的浓厚兴趣. 经过多年的努力，欧拉证明了费马大多数天才的论断，也发现一些错误. 比如说，很容易看出猜想 3 是对的，欧拉证明了猜想 4（即引理 4.1）并且作了推广（引理 4.2）. 我们在下节要介绍欧拉对猜想 1 和 2 的证明. 另一方面，欧拉发现 $2^{2^5}+1$不是素数，即猜想 6 不成立. 只剩下一个猜想没有解决，就是本节一开始叙述的猜想. 欧拉只证明了这个猜想的 $n=3$ 情形，即 $x^3+y^3=z^3$ 没有正整数解.

现在我们介绍费马是如何证明方程 $x^4+y^4=z^4$ 没有正整数解的. 他实际上证明了比这更强的结果.

定理 5.1 方程 $x^4+y^4=z^2$ 没有正整数解.

证 我们用反证法. 如果 $x^4+y^4=z^2$ 存在正整数解，我们在所有正整数解中取（x_0，y_0，z_0），使得 z_0 的值最小. 这时必然 x_0，y_0 和 z_0 两两互素. 比如说若素数 p 是 x_0 和 z_0 的公因子，由 $x_0^4+y_0^4=z_0^2$ 可知 p 也是y_0 的因子. 于

是 $p^4 \mid x_0^4 = y_0^4 = z_0^2$，所以 $p^2 \mid z_0$. 现在 $\left(\frac{x_0}{p}, \frac{y_0}{p}, \frac{z_0}{p^2}\right)$也是 $x^4 + y^4 = z^2$ 的正整数解. 这就与我们事先假定 z_0 的最小性相矛盾. 所以 x_0 和 z_0 互素. 类似可证$(x_0, y_0) = (y_0, z_0) = 1$. 这表明 $x^2 + y^2 = z^2$ 有正整数解$(x, y, z) = (x_0^2, y_0^2, z_0)$，并且 x_0^2, y_0^2, z_0 是两两互素的. 根据第 2 节关于 $x^2 + y^2 = z^2$ 的结果，可知(不妨设 y_0 为偶数)

$$x_0^2 = u^2 - v^2, \quad y_0^2 = 2uv, \quad z_0 = u^2 + v^2,$$

其中$(u, v) = 1$，u 和 v 一奇一偶，并且 $u > v > 0$.

如果 u 为偶数，则 v 为奇数. 从而 $x_0^2 = u^2 - v^2 \equiv -1 \pmod 4$，这不可能. 因此必然 u 为奇数而 v 为偶数. 由$\left(\frac{y}{2}\right)^2 = u \cdot \frac{v}{2}$和$\left(u, \frac{v}{2}\right) = 1$

可知 u 和$\frac{v}{2}$均为正整数的平方

$$u = r^2, \quad \frac{v}{2} = s^2,$$

其中 r, s 为正整数，$(r, s) = 1$ 并且 r 为奇数. 于是

$$x_0^2 = u^2 - v^2 = r^4 - 4s^4.$$

由于$(2s^2, r^2) = 1$，所以$(x, y, z) = (x_0, 2s^2, r^2)$又是$x^2 + y^2 = z^2$的一组基本解，再用定理我们有

$$x_0 = \rho^2 - \sigma^2, \quad 2s^2 = 2\rho\sigma, \quad r^2 = \rho^2 + \sigma^2,$$

其中ρ和σ是互素的正整数，并且$\rho > \sigma > 0$，由于$s^2 = \rho\sigma$，可知ρ和σ都是正整数平方：

$$\rho = f^2, \sigma = g^2,$$

其中f和g是互素的正整数. 于是

$$r^2 = \rho^2 + \sigma^2 = \rho^4 + \sigma^4.$$

这表明$(x, y, z) = (\rho, \sigma, r)$是方程$x^4 + y^4 = z^2$的正整数解. 但是

$$z_0 = u^2 + v^2 = r^4 + 4s^4 > r > 0,$$

这就与z_0的最小性相矛盾. 这一矛盾表明方程$x^4 + y^4 = z^2$不可能有整数解. 证毕.

作为定理5.1的直接推论，便知$x^4 + y^4 = z^4$也不能有正整数解. 费马自己把他的证明方法起了个名字，称为"无穷下降法". 这是因为上面证明的实质是：如果方程$x^4 + y^4 = z^2$有正整数解(x_0, y_0, z_0)，我们可以由此得到另一组正整数解(ρ, σ, r)，其中r是比z_0小的正整数. 再作下去，又可得到另一组正整数解(ρ', σ', r')，使得r'又是比r小的正整数. 这样无穷

地下去是不可能的，因为正整数 $z_0, r, r', \cdots$ 不能无穷地下降．所以只能在一开始的那个正整数解是不可能存在的．

有人猜测费马可能过分估计了无穷下降法的威力．也许他认为用此法可以证明费马猜想的一般情形，但事实远非如此．尽管这样，费马的无穷下降法证明思想在后人的数论研究中仍旧起了重要的作用．例如在20世纪初期，英国数学家莫代尔用此法证明了椭圆曲线的有理点群是有限生成的．

习题：用无穷下降法证明方程 $x^4 + y^2 = z^4$ 没有正整数解．$x^4 + 4y^4 = z^2$ 也没有正整数解．

费马猜想的进一步突破是由欧拉做出的．欧拉研究了并且（肯定或否定地）解决了费马所有的其他猜想，但是对于这里所述的“最后”猜想，欧拉在1770年只能证明对 $n = 3$ 的情形，即证明了方程 $x^3 + y^3 = z^3$ 没有正整数解．又过了半个世纪，法国数学家勒让德和狄里赫利于1825年独立地证明了 $n = 5$ 的情形．1839年，法国另一位数学家拉梅证明了 $n = 7$ 的情形．而费马猜想更重大的进步是在1847年发生的，我们将在第7节介绍．

在这里，我们还想介绍在这个时期一位研究费马猜想的法国女数学家热尔曼和她的传奇性的故事．

索菲·热尔曼(Sophie Germain, 1776～1831)生于1776年4月1日,父亲是一位法国商人.当她还是一位少女时,被一本数学历史书中描写古希腊数学家阿基米德的故事所感动.当罗马入侵士兵来到面前,80岁的阿基米德仍在全神贯注地研究画在沙堆上的几何图形.士兵用长矛刺死他之前,他最后一句话是:不要动我的几何图形! 年轻的热尔曼想:如果一个人能够如此迷恋于比死亡还重要的几何问题,那么数学一定是世界上最迷人的学问.她开始自学微积分和数论,钻研牛顿和欧拉的著作,经常工作到深夜.她对数学的热爱一开始遭到父母的强烈反对,没收她的蜡烛、衣服和取暖的东西,以阻止她学习数学.后来她的执著终于感动了父母,同意并支持她学习数学.

1794年巴黎建立了一所著名的大学:巴黎综合工科学校,用来为法国培养优秀的科学家和数学家,但是只招收男学生.热尔曼冒用一个已经离校男生勒布朗(Le Blanc)的名字偷偷在学校里念书.学校不知真正的勒布朗已经离校,继续给他发讲义和习题.热尔曼使用这些教材并且以勒布朗的名字交作业.两个月之后,数学老师——著名法国数学家拉格朗日发现:一向糟透了的勒布朗近来的作业表现出非凡的数学才华,便主动地要求见这位学生.于是热尔曼暴

露了身份，这使拉格朗日很震惊，并表示愿意做为这位年轻女学生的导师和朋友.

从此之后，热尔曼终于有了一位能激励她前进的老师，可以对他坦诚地展示她的数学才能和抱负.她对于数论抱有最大的兴趣.那时，关于费马猜想只证明了 $n=4$ 和 $n=3$ 这两种情形(当然，由此推出当 n 是 3 或 4 的倍数时，费马猜想也成立).而热尔曼得到了相当一般的结果:

热尔曼定理 设 p 为奇素数，如果 $2p+1$ 也是素数，则方程 $x^p+y^p=z^p$ 没有正整数解 (x,y,z) 使得 $p\nmid xyz$.

例如 $p=3,5,11,13,29,41,53,83$ 和 89 都满足定理所述条件.热尔曼只是对这些素数 p 证明了满足 $p\nmid xyz$ 的正整数解是不存在的，这称为费马猜想的第一种情形.她没有对这些 p 完全证明费马猜想，即没有证明第二种情形:$p\mid xyz$ 的正整数解也不存在.但是在当时这是一个令人吃惊的一般性结果，并且后人沿此方向对热尔曼的工作加以推广.比如勒让德就把它推广成:如果 p 是奇素数，并且在 $4p+1,8p+1,10p+1,14p+1$ 和 $16p+1$ 当中只要有一个也是素数，则当 $n=p$ 时费马猜想的第一种情形成立.由此可得对 100 以内的所有奇素数 p，方程 $x^p+y^p=z^p$ 没有正整数解使得 $p\nmid$

xyz.从热尔曼之后，费马猜想常常对第一和第二种情形分开研究.一般来说，第二种情形更加困难.

热尔曼对费马猜想做了多年研究之后，她想找一位最好的数学家去讨论.她所找的数学家就是她所读过的《算术探究》一书的作者：高斯.高斯当时已是一位杰出的数论学家，但当时他对费马猜想并没有多大兴趣.1816年，巴黎科学院以3000法朗和金质奖章为费马猜想设奖征解.天文学界朋友鼓励高斯参加这项研究活动.但高斯在回信中写道："费马猜想作为一个孤立的命题对我来说几乎没有什么兴趣，因为我可以很容易地写下许多这样的命题，人们既不能证明它们，又不能推翻它们."

至于高斯当时感兴趣的数论问题是什么，我们将在下节介绍其中的一个.现在继续谈热尔曼的故事.热尔曼仍化名勒布朗先生给在德国的高斯小心翼翼地写了第一封信，表示"对于打扰一位天才我深感鲁莽".而高斯尽管对费马猜想兴趣不大，但在回信中热情地鼓励"勒布朗先生"：我很高兴算术找到了像你这样有才能的朋友.后来热尔曼在通信中说出了自己的真实身份，高斯惊喜万分，以致于忘记了自己对费马猜想的消极态度.下面是高斯给热尔曼回信中的一段话.

不知道该怎样向你描述我的钦佩和震惊，当我明白我尊敬的通信者勒布朗先生把自己变成作出了如此辉煌的使我难以相信的范例的卓越人物时.一般而言，对抽象的科学，尤其是对神秘的数论的爱好是非常罕见的.这门高尚科学只对那些有勇气深入其中的人展现其迷人的魅力.而当一位在世俗和偏见的眼光看来一定会遭遇比男子多得多的困难才能通晓这些艰难的研究的女性终于成功地越过种种障碍，洞察其中最令人费解的部分时，那么毫无疑问她一定具有最崇高的勇气、超常的才智和卓越的创造力.事实上，还没有任何东西能以如此令人喜欢和毫不含糊的方式向我证明，这门为我的生活增添了无比欢乐的科学所具有的吸引力绝不是虚构的，如同你的偏爱使它更为荣光一样.

高斯的热情鼓励对于热尔曼的许多工作起了很大的促进作用.1808 年之后，高斯被聘为格丁根大学天文学教授，他的兴趣从数论转到应用数学方面，不再给热尔曼回信.热尔曼也逐渐放弃了纯粹数学研究，改为物理学科，研究“弹性板的振动”.她对费马猜想和弹性板振动的研究荣获法国科学院的金质奖章，成了第一位出席科学院讲座的女性科学家.高斯说服格丁根大学授予她名誉博士学位.但遗憾的是，在格丁根大学授予她这个荣誉之前，她已于 1831

年因乳腺癌去世.她终生未嫁,被誉为"造诣最深并且一生潜心于学术研究的法国女性".

17 世纪和 18 世纪,除了欧拉之外,数论的研究中心在法国.法国是费马猜想的诞生地,也是数论研究和费马猜想早期研究的主要场所. 19 世纪初期在德国出现了伟大的数论学家高斯.高斯对费马猜想本身没有作过重大贡献,但是他创造的数论思想直接影响了费马猜想以后的发展,导致在 19 世纪中期另一位德国数学家库默尔对费马猜想作出了重大贡献.基于此,在我们继续讲述费马猜想之前,打算用一小节讲述高斯对数论的一个贡献,即高斯对二平方和问题的研究和他所采用的新数学工具:(现在被称为)高斯整数环.

6 二平方和问题和高斯整数环

二平方和问题是问:哪些正整数可以表成两个整数的平方和?由 $1=1^2+0^2$, $2=1^2+1^2$, $4=2^2+0^2$, $5=1^2+2^2$ 可知 1,2,4,5 都是二平方和,易知 3,6 和 7 不是二平方和.当 n 是奇素数 p 的特殊情形,费马提出了如下的论断(即上节所述的猜想 1 和 2):如果 $p\equiv3 \pmod 4$,则 p 不能表成二平方和;而若 $p\equiv1 \pmod 4$,则 p 一定可表成二平方和,并且表达方式本质上是惟一的.所谓"本质上惟一"的含义是说:若 $p\equiv1 \pmod 4$,则方程 $p=x^2+y^2$ 有整数解 $(x,y)=(a,b)$,并且它只有 8 组整数解 $(x,y)=(\pm a,\pm b)$ 和 $(\pm b,\pm a)$.

费马的这两个论断是由欧拉于 18 世纪证明的.第一个论断的证明很容易,我们甚至可以证明:任何正整数 $n\equiv3 \pmod 4$, n 均不能表成二平方和.这是因为:每个整数模 4 同余于 1 或

0，所以二平方和 x^2+y^2 模 4 只能同余于 0，1 或 2，不可能同余于 3．费马的第二个论断的证明要困难得多，我们将在后面给出证明．到了 19 世纪初期，德国大数学家高斯在 1801 年所写的书《算术探究》中，完全地解决了二平方和问题，即完全决定出哪些正整数是二平方和，而哪些则不能表成二平方和．高斯在研究二平方和问题中引用了一个新的观点，这种观点对于 19 世纪之后数论的发展起了重要的影响．不仅在后人研究费马猜想时发挥了作用，而且发展出数论的一个重要分支：代数数论．在高斯之后，世界的数论中心从法国转到德国．

高斯的想法看起来很简单，就是引进复数 $i=\sqrt{-1}$ 之后，方程 $n=x^2+y^2$ 变成

$$n=(x+iy)(x-iy).$$

这种变换现在连中学生都会做，可是在 19 世纪初期，那是复数没有被人普遍接受的年代．如果 n 是二平方和，那么 n 可写成复数 $\alpha=x+iy$ 和它的共轭 $\bar{\alpha}=x-iy$ 之乘积，其中 x 和 y 都是整数．于是高斯考虑所有形如 $x+iy$（x，y 是整数）组成的集合，这个集合表示成 $\mathbb{Z}[i]$．于是：正整数 n 是两个整数的平方和，当且仅当 n 是 $\mathbb{Z}[i]$ 中一对共轭复数的乘积．现在把 $\mathbb{Z}[i]$ 中的复数称为**高斯整数**．设 $\alpha=a+ib$ 和 $\beta=c$

$+id$ 均是高斯整数，即 a,b,c,d 都是通常整数，则

$$\alpha \pm \beta = (a \pm c) + i(b \pm d),$$

$$\alpha\beta = (ac - bd) + i(ad + bc)$$

也是高斯整数. 这表明在集合 $\mathbb{Z}[i]$中可进行加减乘运算，所以 $\mathbb{Z}[i]$是一个环，称为**高斯整数环**. 另一方面，

$$\frac{1}{3+4i} = \frac{3-4i}{25} = \frac{3}{25} - \frac{4}{25}i$$

不是高斯整数，所以在 $\mathbb{Z}[i]$中不是永远可以作除法的，非零元素 $3+4i$ 在 $\mathbb{Z}[i]$中不是可逆元素. 我们说过，一个环中可逆元素愈多，作除法就愈灵活. 在通常整数环 $\mathbb{Z}$ 中只有 ± 1 是可逆元素. 在模 m 同余类环 Z_m 中共有 $\varphi(m)$个可逆元素. 现在决定高斯整数环 $\mathbb{Z}[i]$中的可逆元素.

引理 6.1 $\mathbb{Z}[i]$中共有四个可逆元素：± 1 和 $\pm i$.

证 设 $\alpha = a + bi$ 是高斯整数（即 $a, b \in \mathbb{Z}$）. 如果 α 可逆，即存在另一个高斯整数 $\beta = c + di$ $(c, d \in \mathbb{Z})$使得 $\alpha\beta = 1$. 于是 $\bar{\alpha}\,\bar{\beta} = 1$，所以 $\alpha\bar{\alpha} \cdot \beta\bar{\beta} = 1$. 但是 $a^2 + b^2 = \alpha\bar{\alpha}$ 和 $c^2 + d^2 = \beta\bar{\beta}$

都是通常整数，两者乘积是 1，从而必然 $a^2+b^2=c^2+d^2=1$ 于是 $(a,b)=(\pm 1,0)$ 或者 $(0,\pm 1)$. 对于这四种情形，α 只有 ± 1 和 $\pm i$ 这四种可能. 而由 $(\pm 1)^{-1}=\pm 1, i^{-1}=-i, (-i)^{-1}=i$ 知 ± 1 和 $\pm i$ 确实都是 $\mathbb{Z}[i]$ 中可逆元素. 证毕.

高斯引进上述思想的直接好处是下面的结果.

引理 6.2 如果正整数 n 和 m 都是二平方和，则 nm 也是二平方和.

证 设 n 和 m 都是二平方和，则有高斯整数 α 和 β，使得 $n=\alpha\bar{\alpha}$, $m=\beta\bar{\beta}$. 于是 $mn=\alpha\beta\bar{\alpha}\bar{\beta}=\alpha\beta\overline{\alpha\beta}$，其中 $\alpha\beta$ 是高斯整数（因为 $\mathbb{Z}[i]$ 是环）. 这表明 nm 是二平方和. 证毕.

注记 我们可以把证明写的更具体些：设 $n=a^2+b^2$, $m=c^2+d^2$，我们可以取 $\alpha=a+bi$, $\beta=c+di$，于是 $\alpha\beta=(ac-bd)+i(ad+bc)$. 所以 nm 可以写成二平方和 $(ac-bd)^2+(ad+bc)^2$. 当然，直接验证恒等式

$$(a^2+b^2)(c^2+d^2)=(ac-bd)^2+(ad+bc)^2$$

并不困难，但是高斯给出它的道理和发现的途径.

整数 $n\geqslant 2$ 可以表成一些素数的乘积. 如果每个素因子都是二平方和，那么由引理 6.2

便知 n 是二平方和.所以很自然地我们先考虑哪些素数可以表成二平方和.这就使我们回到费马的两个论断上来.我们已经知道 $2=1^2+1^2$,即 2 是二平方和.对于奇素数 p,当 $p\equiv 3 \pmod 4$时 p 不是二平方和,现在证明费马对 $p\equiv 1 \pmod 4$情形作出的论断.

引理 6.3(欧拉) 设 p 是被 4 除余 1 的素数,则 p 是二平方和,并且方程 $p=x^2+y^2$ 恰有八组整数解$(x,y)=(\pm a,\pm b)$和$(\pm b,\pm a)$.

证 为了证明 p 是二平方和,我们首先证明存在整数 m,$1\leqslant m\leqslant p-1$,使得 mp 是二平方和.然后用费马的"无穷下降法";如果 $m\geqslant 2$,则必有比 m 小的正整数 m',使得 $m'p$ 也是平方和.如果 m' 仍旧大于 1,则继续这样做下去,所以 m' 一定下降到 1,即 p 一定是二平方和.

我们先做第一步,在上节的例 2 中我们证明了:若 p 是模 4 余 1 的素数,则存在整数 a,使得 $a^2\equiv -1 \pmod p$.由于$\left\{-\frac{p-1}{2},-\frac{p-1}{2}+1,\cdots,-1,0,1,\cdots,\frac{p-1}{2}\right\}$是模 p 的完系,所以 a 模 p 必同余于某个整数 b,使得 $|b|\leqslant\frac{p-1}{2}$.这时 $b^2\equiv a^2\equiv -1 \pmod p$.于是 b^2+1^2

$=mp$，其中 m 为正整数．但是 $mp=b^2+1\leqslant\left(\frac{p-1}{2}\right)^2+1<p^2$，所以 $m<p$．因此我们有 $1\leqslant m\leqslant p-1$，而 mp 是二平方和．

现在做第二步，即若 $2\leqslant m\leqslant p-1$，$mp$ 是二平方和，我们再找出比 m 小的正整数 m'，使得 $m'p$ 也是二平方和．由于 mp 是二平方和，$mp=x^2+y^2$，其中 x 和 y 为整数．于是 $x^2+y^2\equiv 0\pmod m$．当 m 为奇数时，$\left\{-\frac{m-1}{2},-\frac{m-1}{2}+1,\cdots,-1,0,1,\cdots,\frac{m-1}{2}\right\}$ 是模 m 的完系．而当 m 为偶数时，$\left\{-\frac{m}{2},-\frac{m}{2}+1,\cdots,-1,0,1,\cdots,\frac{m}{2}-1\right\}$ 是模 m 的完系．所以每个整数模 m 都同余于绝对值不超过 $\frac{m}{2}$ 的整数，即存在整数 x' 和 y'，$|x'|\leqslant\frac{m}{2}$，$|y'|\leqslant\frac{m}{2}$，使得 $x\equiv x'\pmod m$，$y\equiv y'\pmod m$．于是 $x'^2+y'^2\equiv x^2+y^2\equiv 0\pmod m$．于是 $x'^2+y'^2=m'm$，其中 m' 为非负整数．我们现在证明 $1\leqslant m'<m$．由 $m'm=x'^2+y'^2\leqslant\left(\frac{m}{2}\right)^2+\left(\frac{m}{2}\right)^2=\frac{m^2}{2}$ 可知 $m'\leqslant\frac{m}{2}<m$．另一方面，若 $m'=0$，则

$x'^2 + y'^2 = 0$，于是 $x' = y' = 0$. 从而 x 和 y 均为 m 的倍数（因为 $x \equiv x' \pmod{m}$，$y \equiv y' \pmod{m}$）. 于是 $mp = x^2 + y^2 \equiv 0 \pmod{m^2}$，从而 $m^2 \mid mp$，即 $m \mid p$. 但这与假设 $2 \leqslant m \leqslant p - 1$ 相矛盾. 这就表明 $m' \geqslant 1$. 现在有

$$
\begin{aligned}
m'm^2 p &= (x^2 + y^2)(x'^2 + y'^2) \\
&= (xx' + yy')^2 + (xy' - x'y)^2, \quad (6.1)
\end{aligned}
$$

但是

$$
xx' + yy' \equiv x^2 + y^2 \equiv 0 \pmod{m},
$$

$$
xy' - x'y \equiv xy - xy \equiv 0 \pmod{m}.
$$

可知 $A = \frac{1}{m}(xx' + yy')$ 和 $B = \frac{1}{m}(xy' - x'y)$ 均为整数，由(6.1)式知 $A^2 + B^2 = m'p$. 于是我们找到了整数 m'，$1 \leqslant m' < m$，使得 $m'p$ 为二平方和. 综合上述，采用无穷下降法即知 p 为二平方和.

设 $p = a^2 + b^2$，其中 a 和 b 为整数. 易知 a 和 b 均不为零，并且 $|a| \neq |b|$. 所以 $p = x^2 + y^2$ 的一组解 $(x, y) = (a, b)$ 可得出八组不同的解 $(x, y) = (\pm a, \pm b)$ 和 $(\pm b, \pm a)$. 我们再证明方程 $p = x^2 + y^2$ 只有这八组整数解. 如果设 a 和 b 均为正整数，我们只需证此方程只有两组正整数解 $(x, y) = (a, b)$ 和 (b, a). 换句话

说，如果$(x,y)=(A,B)$又是一组正整数解，我们只需证明 $a=A$ 或者 $a=B$. 由于

$$p^2=(a^2+b^2)(A^2+B^2)$$
$$=(aA+bB)^2+(aB-bA)^2, \quad (6.2)$$
$$p^2=(a^2+b^2)(A^2+B^2)$$
$$=(aA-bB)^2+(aB+bA)^2, \quad (6.3)$$

并且

$$(aA+bB)(aA-bB)$$
$$=A^2(a^2+b^2)-b^2(A^2+B^2)$$
$$=A^2p-b^2p\equiv 0 \pmod p,$$

可知 $p \mid aA+bB$ 或者 $p \mid aA-bB$. 如果 $p \mid aA+bB$，由于 $aA+bB$ 为正整数和(6.2)式可知必然 $aA+bB=p$，$aB-bA=0$. 于是 $\frac{a^2}{A^2}=\frac{b^2}{B^2}=\frac{a^2+b^2}{A^2+B^2}=\frac{p}{p}=1$，从而 $a=A$. 如果 $p \mid aA-bB$，则同样地由(6.3)式可知 $p \mid aB+bA$. 于是 $aB+bA=p$，$aA-bB=0$. 所以 $\frac{a^2}{B^2}=\frac{b^2}{A^2}=\frac{a^2+b^2}{A^2+B^2}=1$，即 $a=B$. 这就完全证明了引理.

接下来我们再看高斯如何利用欧拉的上述结果完全解决二平方和问题.我们把每个整数的平方称为平方数.每个正整数都存在惟一的最大平方因子.例如 $6000=2^4\cdot3\cdot5^3=(2^2\cdot5)^2\cdot3\cdot5=20^2\cdot15$,因此 20^2 就是 6000 的最大平方因子.一般地,每个正整数 n 可以表示成 $m^2\cdot m'$,其中 m 为正整数,而 $m'=1$ 或者 m' 是不同素数的乘积,这时 m^2 就是 n 的最大平方因子,我们称 m' 为 n 的**无平方因子部分**.

定理 6.4(高斯) 设 n 为正整数,$n=m^2m'$,其中 m' 为 n 的无平方因子部分,则 n 是二平方和的充分必要条件是 $m'=1$ 或者 m' 的所有素因子模 4 均不同余于 3.

证 充分性是容易的:如果 $m'=1$,则 $n=m^2=m^2+0^2$ 是二平方和.如果 $m'=p_1\cdots p_t$,其中每个素因子 p_i 模 4 均不同余于 3,即 $p_i=2$ 或者 $p_i\equiv1\pmod 4$,由欧拉结果(引理 6.3)知所有的 p_i 都是二平方和.再由引理 6.2 知 $m'=p_1\cdots p_t$ 是二平方和.于是 $n=m^2m'$ 也是二平方和.

现在对 n 用归纳法证明必要性.即我们归纳证明如下的命题:

若正整数 n 是二平方和,则 n 的无平方因子部分 m' 或者为 1,或者其素因子模 4 均不同余于 3.

当 $n=1$ 时命题显然成立.现在设 $n\geqslant 2$ 并且命题对小于 n 的正整数均成立.我们来证命题对 n 成立.由假设 n 是平方和:$n=x^2+y^2$,其中 x 和 y 是整数.如果 n 的所有素因子均是平方和,则 m' 的素因子也是如此,即均模 4 不同余 3,从而命题对 n 成立.下设 n 有素因子 $p\equiv 3 \pmod 4$.于是 $x^2+y^2\equiv 0 \pmod p$.如果 x 不被 p 除尽,则 y 也不被 p 除尽.我们有 $\left(\frac{x}{y}\right)^2 \equiv -1 \pmod p$.但是在 $p\equiv 3 \pmod 4$ 时这是不可能的(第 4 节例 2),所以必然 $p\mid x$,于是 $p\mid n-x^2=y^2$,即 $p\mid y$.所以 $p^2\mid x^2+y^2=n$.这表明 $x'=\frac{x}{p}$,$y'=\frac{y}{p}$ 是整数,并且 $n'=\frac{n}{p^2}$ 是正整数,由 $n'=x'^2+y'^2$ 知 n' 为二平方和.但是 $n'<n$,所以由归纳假设知 n' 的无平方因子部分的素因子模 4 均不同余 3.但是 $n=p^2n'$ 和 n' 有相同的无平方因子部分.这就表明 n 的无平方因子部分的素因子模 4 均不同余 3.所以命题对 n 成立.证毕.

例 试问 1999 和 2000 是否为平方和.

解 由于 $1999\equiv 3 \pmod 4$,可知 1999 不是二平方和.另一方面,$2000=2^4\times 5^3=(2^2\cdot 5)^2\cdot 5$,无平方因子部分为 5.由定理 6.4 知 2000 是平方和.事实上,$5=1^2+2^2$,于是 $2000=400\times 5=20^2+40^2$.

实际上,2000还有其他方法表示成二平方和,例如 $2000=44^2+8^2$.这就产生出进一步的问题:正整数 n 一共有多少方式可以表成二平方和? 即方程 $x^2+y^2=n$ 一共有多少组整数解?(由于 x 和 y 的绝对值均不超过$\sqrt{n}$,所以整数解只有有限多组.)高斯解决了这个问题.高斯的方法是仔细研究环 $\mathbb{Z}[i]$,发现这个环有和整数环 $\mathbb{Z}$ 相类似的惟一因子分解定理.我们在下节将会看到,高斯对于 $\mathbb{Z}[i]$中惟一因子分解性质的研究对后人研究费马猜想有重要的启示.

高斯仿照2000年前欧几里得对于整数环的方式来建立环 $\mathbb{Z}[i]$中的高斯整数分解理论.首先可以完全同样地定义整除性:设 α 和 β 是两个高斯整数,$\alpha\neq 0$.称 α 整除 β,是指 β/α 为高斯整数,即存在高斯整数 γ,使得 $\beta=\alpha\gamma$.这时,称 α 是 β 的因子.接下来要定义什么是高斯素数.它们应当是不能再分解成两个高斯整数之乘积的那些高斯整数.在讨论正整数分解时,我们要排除掉平凡的分解 $n=1\cdot n=n\cdot 1$.现在我们需要排除掉更多的平凡分解,因为高斯整数环 $\mathbb{Z}[i]$中有四个可逆元素 $U=\{\pm 1,\pm i\}$(引理6.1),对每个可逆元素 $\varepsilon\in U$,ε^{-1}也是高斯整数,从而每个高斯整数 α 都可写成 ε^{-1}与$(\alpha\varepsilon)$之乘积.这样的分解都称为是平凡

的.所以合理的高斯素数应当定义成如下的形式:

定义 6.5 非零高斯整数 π 称为**高斯素数**,是指 π 只有平凡的分解,即若 π 是高斯整数 α 和 β 的乘积,则 α 和 β 当中必有一个为可逆元素.

为了给出高斯素数的另一种定义方式,我们引进一个符号.对每个高斯整数 $\alpha = a + bi$(其中 $a, b \in \mathbb{Z}$),我们把 α 的绝对值的平方

$$N(\alpha) = |\alpha|^2 = a^2 + b^2$$

称为 α 的**范数**. $N(\alpha)$是非负整数,并且当 $\alpha \neq 0$ 时,$N(\alpha)$是正整数.进而,不难证明:α 是可逆元素当且仅当 $N(\alpha) = 1$.如果 $\alpha = \beta\gamma$,则 $N(\alpha) = N(\beta)N(\gamma)$.如果 β 和 γ 都不是可逆元素,则 $N(\beta)$和 $N(\gamma)$均大于 1,这也相当于 $N(\beta)$ 和 $N(\gamma)$均小于 $N(\alpha)$.因此我们可以说:

非零高斯整数 α 是高斯素数,是指它不能写成高斯整数 β 和 γ 的乘积,使得 $N(\beta)$和 $N(\gamma)$都小于$N(\alpha)$.

从高斯素数的这种刻画方式,我们能够判别:若 α 为高斯整数并且 $N(\alpha) = p$ 是通常的素数,则 α 必为高斯素数.这是因为:如果 $\alpha = \beta\gamma$,其中 $\beta, \gamma \in \mathbb{Z}[i]$,则 $N(\beta)N(\gamma) = N(\alpha) = p$.但是 $N(\beta)$和 $N(\gamma)$都是正整数,所以必

然有一个为 1，即 β 和 γ 必然有一个为可逆元（另一个的范数为 $N(\alpha)$）. 这表明 α 是高斯素数. 例如 $N(1+i)=2$，$N(1+2i)=5$，所以 $1+i$ 和 $1+2i$ 都是高斯素数.

我们也可以看出：若 α 是非零的高斯整数并且不是可逆元，则 α 一定是有限个高斯素数之积. 因若 α 是高斯素数则分解完毕. 若不然，则 $\alpha=\beta\gamma$，其中 $\beta,\gamma\in\mathbb{Z}[i]$，并且 $N(\beta)$ 和 $N(\gamma)$ 都比 $N(\alpha)$ 小. 将 β 和 γ 再如此作下去，由于范数不断变小，所以必在有限步之后停止. 这时，α 就表示成有限个高斯素数的乘积.

接下来的问题是，我们应当期望高斯整数的这种分解有怎样的惟一性？首先，像在 $\mathbb{Z}$ 中的情形一样，每个分解式的因子可以任意调换次序. 其次还要考虑可逆元素的影响，因为若 $\alpha=\beta\gamma$，则对每个可逆元素 ε，$\alpha=(\varepsilon\beta)\cdot(\gamma\varepsilon^{-1})$. 我们应当把这两种分解也看成本质上是一样的. 这就引导出如下的概念.

定义 6.6 两个非零高斯整数 α 和 β 称为是**相伴**的（表示成 $\alpha\sim\beta$），是指存在可逆元 ε，使得 $\alpha=\beta\varepsilon$.

不难验证，非零高斯整数的相伴关系具有自反性、对称性和传递性（这原因是由于可逆元素集合 $U=\{\pm 1,\pm i\}$ 是乘法群，即若 ε,λ 为可逆元素，则 ε^{-1} 和 $\varepsilon\lambda$ 也是可逆元素）. 所以相

伴关系是等价关系，非零高斯整数由此分成许多相伴类，每类有 4 个高斯整数 $\{\pm\alpha, \pm i\alpha\}$. 易知若 $\alpha \sim \beta$，则 $N(\alpha) = N(\beta)$. 由此不难证明：若 α 是高斯素数，$\beta \sim \alpha$，则 β 也是高斯素数. 如果要把分解 $\alpha = \beta\gamma$ 和 $\alpha = (\varepsilon\beta)(\gamma\varepsilon^{-1})(\varepsilon \in U)$看成本质上是一样的，就需要把相伴元($\beta$ 和 $\varepsilon\beta$，γ 和 $\gamma\varepsilon^{-1}$)不加以区别. 高斯证明了：当高斯整数分解成高斯素数乘积时，如果不考虑因子的次序和相伴性，则分解式是惟一的. 也就是说，高斯整数环 $\mathbb{Z}[i]$具有如下的惟一分解特性.

定理 6.7(高斯) 设 α 是 $\mathbb{Z}[i]$中非零元素并且不是可逆元素. 则

(1)(存在性) α 可以表示成有限个高斯素数的乘积.

(2)(惟一性) 如果 $\alpha = \pi_1\pi_2\cdots\pi_s = \lambda_1\lambda_2\cdots\lambda_t$ 是 α 的两种分解，其中 $\pi_i(1 \leqslant i \leqslant s)$和 $\lambda_j(1 \leqslant j \leqslant t)$都是高斯素数，则 $t = s$，并且适当调整 $\pi_1, \pi_2, \cdots, \pi_s$ 的次序，可以使 $\pi_i \sim \lambda_i(1 \leqslant i \leqslant t)$.

我们已经在前面证明了分解的存在性. 限于篇幅，在这里略去惟一性的证明，只是告诉大家，高斯对于 $\mathbb{Z}[i]$的分解惟一性的证明本质上是仿照欧几里得对 $\mathbb{Z}$ 中分解惟一性所用的方法.

高斯利用定理 6.7 完全决定出全部高斯素数.

定理 6.8 设 p 是(通常的)素数.

(1) 对于 $p=2$, $2=(1+i)(1-i)$, 并且 $1+i$ 和 $1-i$ 是相伴的高斯素数.

(2) 若 $p\equiv1\pmod 4$, 则 $p=\pi\bar{\pi}$, 其中 π 和 $\bar{\pi}$ 是不相伴的高斯素数.

(3) 若 $p\equiv3\pmod 4$, 则 p 是高斯素数.

(4) 由上述方式给出的高斯素数和它们的相伴元便是全部高斯素数.

证 (1) 由 $N(1\pm i)=2$ 为素数, 可知 $1\pm i$ 均是高斯素数. 由 $\frac{1+i}{1-i}=i$ 是可逆元素, 可知 $1+i\sim1-i$.

(2) 当 $p\equiv1\pmod 4$ 时 p 是二平方和, 即 $p=a^2+b^2=(a+bi)(a-bi)$, 其中 $a,b\in\mathbb{Z}$. 由于 $N(a\pm bi)=p$ 可知 $a\pm bi$ 均是高斯素数. 请大家证明 $a+bi$ 和 $a-bi$ 是不相伴的 $\left(\text{事实上}, \frac{a+bi}{a-bi}\text{不属于}\mathbb{Z}[i], \text{所以更不可能为}\mathbb{Z}[i]\text{中可逆元素}.\right)$

(3) 设 $p\equiv3\pmod 4$. 如果 p 不是高斯素数, 则 $p=\alpha\beta$, 其中 α 和 β 是高斯整数, 并且 $N(\alpha)$ 和 $N(\beta)$ 均小于 $N(p)=p^2$. 因此必然 $N(\alpha)=p$. 记 $\alpha=a+bi\ (a,b\in\mathbb{Z})$, 则 $p=$

$N(\alpha)=a^2+b^2$，但是 $p\equiv 3\pmod 4$ 时 p 不是二平方和．这个矛盾表明 p 一定是高斯素数．

(4) 对于每个高斯素数 π，$N(\pi)=\pi\bar{\pi}=n$ 是正整数．将 n 先在 中分解成一些素数的乘积：$n=p_1p_2\cdots p_s$，再将每个素数 p_i 按着本定理的(1)，(2)，(3)分解成高斯素数乘积．于是得到 $n=\lambda_1\lambda_2\cdots\lambda_t$．从而 $\pi\bar{\pi}=\lambda_1\lambda_2\cdots\lambda_t$．利用分解的惟一性可知 π 必然与某个 λ_i 相伴，而 λ_i 是从 n 的某个素数因子 p_j 分解得来的．这就完成了证明．

利用 $[i]$ 中素因子分解的存在性，高斯给出二平方和结果(定理 6.4)的另一个证明．设 $n=m^2m'$，其中 m' 是正整数 n 的无平方因子部分．如果 m' 没有素因子 $p\equiv 3\pmod 4$，可以像前面那样容易证明 n 为二平方和．反之，若 $n=a^2+b^2(a,b\in\ )$，令 $\alpha=a+bi$，则 $n=\alpha\bar{\alpha}=N(\alpha)$．现在把高斯整数 α 分解成高斯素数之乘积：$\alpha=\pi_1\pi_2\cdots\pi_s$，则 $n=N(\alpha)=N(\pi_1)N(\pi_2)\cdots N(\pi_s)$．每个素数 $p\equiv 3\pmod 4$ 都是高斯素数．如果 π_i 与 p 相伴，则 $N(\pi_i)=N(p)=p^2$．如果 π_i 不与 p 相伴，由定理 6.8 可知 π_i 是由其他素数 $p'(\neq p)$ 分解得来的，所以 $N(\pi_i)=p'$．如果 $\pi_1,\cdots,\pi_s$ 当中恰好有 t 个与 p 相伴，则 $n=N(\pi_1)N(\pi_2)\cdots N(\pi_s)$的右边恰

好有 p^{2t}，其余因子是 p 以外的素数.这就表明每个素数 $p\equiv 3(\bmod 4)$ 在 n 的分解式中都有偶数 $(2t)$ 个，因此都不包含在 n 的无平方因子部分 m' 之中.这就证明了定理 6.4.

最后我们再谈一下高斯如何用 $[i]$ 中分解的惟一性来计算方程 $n=x^2+y^2$ 的整数解个数 $f(n)$.我们知道，$n=a^2+b^2(a,b\in\ )$ 相当于高斯整数 $\alpha=a+bi$ 的范数为 n.所以方程 $n=x^2+y^2$ 的整数解个数 $f(n)$ 就是范数为 n 的高斯整数的个数.但是若 α 的范数为 n，则相伴类 $\{\pm\alpha,\pm i\alpha\}$ 中每个高斯整数的范数都为 n.如果以 $F(n)$ 表示范数为 n 的高斯整数形成的相伴类数，则 $f(n)=4F(n)$.

定理 6.9(高斯) (1) 若 n 和 m 是互素的正整数，则 $F(nm)=F(n)F(m)$.

(2) 设正整数 $n=2^{a_0}p_1^{a_1}\cdots p_s^{a_s}q_1^{b_1}\cdots q_t^{b_t}$，其中 $p_1,\cdots,p_s,q_1,\cdots,q_t$ 是不同的素数，$p_i\equiv 1(\bmod 4)(1\leqslant i\leqslant s)$，$q_j\equiv 3(\bmod 4)(1\leqslant j\leqslant t)$，而 $a_0,\cdots,a_s,b_1,\cdots,b_t$ 是非负整数.则当 $b_1,\cdots,b_t$ 至少有一个为奇数时，$f(n)=0$.而当 $b_1,\cdots,b_t$ 都是偶数时，$f(n)=4(a_1+1)(a_2+1)\cdots(a_s+1)$.

证 (1) 范数为 1 的只有相伴类 $\{\pm 1,\pm i\}$ 中的高斯整数，因此 $F(1)=1$.由此可知当 $m=1$ 或 $n=1$ 时 $F(mn)=F(m)F(n)$ 成立.

以下设 $n, m \geqslant 2$. 我们先证明如下两个论断：

(A) 若 α 是范数为 mn 的高斯整数，则存在范数分别为 n 和 m 的高斯整数 β 和 γ，使得 $\alpha = \beta\gamma$.

(B) 如果又有范数分别为 n 和 m 的高斯整数 β' 和 γ' 使得 $\alpha = \beta'\gamma'$，则 $\beta \sim \beta'$，$\gamma \sim \gamma'$.

假设 $\alpha = \pi_1 \cdots \pi_s \pi'_1 \cdots \pi'_t$，其中 π_i 和 π'_j 均为高斯素数. 则 $nm = N(\alpha) = N(\pi_1) \cdots N(\pi_s) N(\pi'_1) \cdots N(\pi'_t)$. 由于 $N(\pi_i)$ 或 $N(\pi'_j)$ 都是素数或素数的平方，而 n 和 m 互素，可知 $N(\pi_i)$（或 $N(\pi'_j)$）恰好除尽 n 和 m 当中的一个. 我们不妨设 $N(\pi_i)$ $(1 \leqslant i \leqslant s)$ 均是 n 的因子，$N(\pi'_j)$ $(1 \leqslant j \leqslant t)$ 均是 m 的因子. 令 $\beta = \pi_1 \cdots \pi_s$，$\gamma = \pi'_1 \cdots \pi'_j$. 那么 $\alpha = \beta\gamma$，$N(\beta)N(\gamma) = mn$. 由于 $N(\beta)$ 与 m 互素可知 $N(\beta) \mid n$，类似可知 $N(\gamma) \mid m$. 再由 $N(\beta)N(\gamma) = mn$ 可知 $N(\beta) = n$，$N(\gamma) = m$. 这就证明了论断(A).

再设 $\alpha = \beta'\gamma'$，$N(\beta') = n$，$N(\gamma') = m$，则 $\beta\gamma = \beta'\gamma'$. 由于 $\beta\bar{\beta} = n$ 和 $\gamma'\bar{\gamma}' = m$ 互素，可知 β 的高斯素因子都不是 γ' 的高斯素因子. 由此可知 $\beta \mid \beta'$. 同样地，由 $\beta'\bar{\beta}' = n$ 和 $\gamma\bar{\gamma} = m$ 互素以及 $\beta\gamma = \beta'\gamma'$ 可知 $\beta' \mid \beta$. 这就表明 $\beta \sim \beta'$，从而 $\gamma \sim \gamma'$. 这就证明了论断(B).

$N(\alpha)=n$ 的 α 组成 $F(n)$ 个相伴类，$N(\beta)=m$ 的 β 组成 $F(m)$ 个相伴类．由论断(A)和(B)可知，若将前 $F(n)$ 个相伴类分别取出一组代表元 $\alpha_1,\cdots,\alpha_t(t=F(n))$，后 $F(m)$ 个相伴类也分别取出一组代表元 $\beta_1,\cdots,\beta_s(s=F(m))$，则 $\alpha_i\beta_j(1\leqslant i\leqslant F(n),1\leqslant j\leqslant F(m))$ 恰好是范数为 nm 的元素组成的 $F(nm)$ 个相伴类的完全代表系．这就证明了 $F(nm)=F(n)F(m)$．

(2) 由(1)可知 $F(n)=F(2^{a_0})F(p_1^{a_1})\cdots F(p_s^{a_s})F(q_1^{b_1})\cdots F(q_t^{b_t})$．若 $b_1,\cdots,b_t$ 当中某个 b_j 为奇数时，则因 $q_j\equiv 3(\bmod 4)$，由定理 6.4 可知 n 不是二平方和，即 $F(n)=0$．以下设 $b_j(1\leqslant j\leqslant t)$ 均为偶数．记 $b_j=2c_j(1\leqslant j\leqslant t)$，其中 c_j 是非负整数．若 $N(\alpha)=q_j^{2c_j}$，则 $\alpha\bar{\alpha}=q_j^{2c_j}$．由于 q_j 是高斯素数，由分解的惟一性可知 α 和 $\bar{\alpha}$ 的每个高斯素因子均与 q_j 相伴．但是 α 和 $\bar{\alpha}$ 应当是同样多个高斯素数的乘积(若 $\alpha=\pi_1\cdots\pi_l$，则 $\bar{\alpha}=\bar{\pi}_1\cdots\bar{\pi}_l$)，所以 α 一定相伴于 $q_j^{c_j}$．这就表明 $N(\alpha)=q_j^{2c_j}$ 的 α 只有一个相伴类，即 $F(q_j^{2c_j})=1(1\leqslant j\leqslant t)$．进而，$\pi=1+i$ 是高斯素数并且 $2=(1+i)(1-i)\sim\pi^2$．如果 $N(\alpha)=2^{a_0}$，则 $\alpha\bar{\alpha}=2^{a_0}\sim\pi^{2a_0}$，可知 $\alpha\sim\pi^{a_0}$，所以也有 $F(2^{a_0})=1$．最后计算 $F(p_i^{a_i})(1\leqslant i\leqslant s)$．由 p_i

$\equiv 1 \pmod 4$ 可知 $p_i = \pi_i\overline{\pi}_i$，其中 $N(\pi_i) = N(\overline{\pi}_i) = p_i$ 并且 π_i 和 $\overline{\pi}_i$ 不相伴. 如果 $N(\alpha) = p_i^{a_i}$，则 $\alpha\overline{\alpha} = \pi_i^{a_i}\overline{\pi}_i^{a_i}$. 满足这个等式并且彼此不相伴的 α 只有 $\pi_i^{a_i}, \pi_i^{a_i-1}\overline{\pi}_i, \cdots, \pi_i\overline{\pi}_i^{a_i-1}$ 和 $\overline{\pi}_i^{a_i}$ 这 a_i+1 个数. 所以 $F(p_i^{a_i}) = a_i + 1 (1 \leqslant i \leqslant s)$. 综合上述便知 $F(n) = (a_1+1)\cdots(a_s+1)$ 和 $f(n) = 4(a_1+1)\cdots(a_s+1)$.

例 求 $x^2 + y^2 = 2000$ 的全部整数解.

解 $2000 = 2^4 \cdot 5^3$. $f(2000) = 4F(2000) = 4\cdot(3+1) = 16$. 由于 $F(2^4) = 1$，范数为 2^4 的高斯整数必相伴于 $2^2 = 4$. 由于 $F(5^3) = 4$，$5^3 = (1+2i)^3(1-2i)^3$，所以范数为 5^3 的高斯整数必相伴于 $\alpha_1 = (1+2i)^3 = -11-2i$，$\alpha_2 = (1+2i)^2(1-2i) = 5+10i$，$\alpha_3 = (1+2i)(1-2i)^2 = 5-10i$ 和 $\alpha_4 = (1-2i)^3 = -11+2i$ 当中的一个. 所以范数为 2000 的高斯整数必相伴于 $4\alpha_i (1 \leqslant i \leqslant 4)$ 当中的一个. 每个 $4\alpha_i$ 又相伴于 4 个高斯整数. 由此给出 16 个范数为 2000 的高斯整数. 从而给出方程 $2000 = x^2 + y^2$ 的全部整数解. 它们是 $(x, y) = (\pm 44, \pm 8), (\pm 8, \pm 44), (\pm 20, \pm 40)$ 和 $(\pm 40, \pm 20)$.

我们花了较多篇幅讲述高斯整数环 $\mathbb{Z}[i]$，目的是想描述高斯研究数论问题的一个思想. 他把本来属于环 $\mathbb{Z}$ 中的二平方和问题放在更大

的环 $[i]$中去研究，发现了环 $[i]$的惟一因子分解特性并由此解决了二平方和问题．这种方法对后来的数论研究起很大的影响，而且发展出数论的一个分支：代数数论．费马猜想的重要一步便是沿这种思想得出的．

7 库默尔的贡献

现在继续讲费马猜想.费马本人对于 $n=4$ 的情形证明了 $x^n+y^n=z^n$ 没有正整数解.假设 $n=ms$,其中 m 和 s 均是正整数,如果方程 $x^n+y^n=z^n$ 有正整数解(a,b,c),由于$(a^s)^m+(b^s)^m=a^n+b^n=c^n=(c^s)^m$,可知$(a^s,b^s,c^s)$是方程 $x^m+y^m=z^m$ 的正整数解.或者反过来说,若费马猜想对 m 成立,即 $x^m+y^m=z^m$ 没有正整数解,则费马猜想对 m 的每个正倍数 n 也成立.由于每个正整数 $n\geqslant3$ 必有 4 或者奇素数为因子,所以若对每个奇素数 p 都能证明费马猜想成立,我们便知对任意 $n\geqslant3$ 费马猜想都成立.于是只需要考虑 $n=p$ 为奇素数的情形就可以了.

1770 年,欧拉对于 $n=3$ 的情形证明了费马猜想,他引进

$$\omega=e^{\frac{2\pi i}{3}}=\cos120^\circ+i\sin120^\circ$$

$$=\frac{1}{2}(-1+\sqrt{3}i).$$

由于 $\omega^3=1,\omega\neq 1$，所以 ω 是多项式 $\frac{x^3-1}{x-1}=x^2+x+1$的根. 由此也可算出 $\omega=\frac{1}{2}(-1+\sqrt{3}i)$，另一根为 $\bar{\omega}=\omega^2=\frac{1}{2}(-1-\sqrt{3}i)$. 引入 ω 之后，方程 $x^3+y^3=z^3$ 就变成

$$y^3=z^3-x^3=(z-x)(z-\omega x)(z-\omega^2 x). \tag{7.1}$$

于是变为乘积的形式（就像高斯引入 i 之后把 x^2+y^2 变为乘积 $(x+iy)(x-iy)$ 一样）. 以 $[\omega]$表示形如 $a+b\omega$ $(a,b\in$) 的复数组成的集合，这样的数相加减显然还是这样的数，因为 ω 是 x^2+x+1 的根，即 $\omega^2=-\omega-1$. 于是对这样的数 $a+b\omega$ 和 $c+d\omega$，乘积

$$\begin{aligned}(a+b\omega)(c+d\omega)&=ac+(ad+bc)\omega+bd\omega^2\\&=(ac-bd)+(ad+bc-bd)\omega\end{aligned}$$

仍是这样的数，所以 $[\omega]$也是一个环. 如果 $x^3+y^3=z^3$有正整数解 (x,y,z)，则(7.1)式的两边就是同一个数 z^3-x^3 在环 $[\omega]$中的两个乘积表达式 y^3 和 $(z-x)(z-\omega x)(z-\omega^2 x)$，欧拉对于环 $[\omega]$中的因子分解特性做了仔细的讨论，通过比较复杂的计算，证明了当 x,y,z 均是正整数时，公式(7.1)不能成立，所

以费马猜想对 $n=3$ 成立.详情可参见 H.M. Edwards 的书"Fermat's Last Theorem"(《费马最后定理》)的第二章(书的作者在这里还指出欧拉的证明有一处不够完善).

1825 年,法国数学家勒让德(1752～1833, Legendre)在他 73 岁的高龄第一个证明了费马猜想对于 $n=5$ 成立,其推导更为复杂,而年青的狄里赫利(Dirichlet)在同一年也作了同样工作.1839 年法国另一位数学家拉梅(Lamé)证明了 $n=7$ 的情形.

1847 年,是数论史上值得纪念的一年.在这一年的 3 月 1 日法国科学院会议上,拉梅宣布他完全证明了费马猜想.他介绍了所用的方法,即对每个奇素数 p,令

$$\zeta_p = e^{\frac{2\pi i}{p}} = \cos\frac{360^\circ}{p} + \left(\sin\frac{360^\circ}{p}\right)i.$$

由于 $\zeta_p^p=1$, $\zeta_p\neq 1$,可知 ζ_p 是$\frac{x^p-1}{x-1}=x^{p-1}+x^{p-2}+\cdots+x+1$的根.如果 $x^p+y^p=z^p$ 有正整数解(x,y,z),则

$$\begin{aligned} y^p &= z^p - x^p \\ &= (z-x)(z-\zeta_p x)(z-\zeta_p^2 x)\cdots(z-\zeta_p^{p-1}x). \end{aligned} \tag{7.2}$$

如果用 $\mathbb{Z}[\zeta_p]$ 表示形如 $a_0 + a_1\zeta_p + \cdots + a_{p-2}\zeta_p^{p-2}$ $(a_0,\cdots,a_{p-1} \in \mathbb{Z})$ 的复数组成的集合,像欧拉对 $p=3$ 的特殊情形所做的那样, $\mathbb{Z}[\zeta_p]$ 中可以进行加减乘运算,用现在的语言来说, $\mathbb{Z}[\zeta_p]$ 是一个环.拉梅说,(7.2)式右边的 p 个因子是彼此互素的,但是(7.2)式左边为 p 次方,所以每个因子 $(z-\zeta_p^i x)$ 也都应当是 p 次方,即 $z-\zeta_p^i x=\alpha_i^p\ (0\leqslant i\leqslant p-1)$,其中 $\alpha_i \in \mathbb{Z}[\zeta]$.然后拉梅说这是不可能的,于是证明了费马猜想.拉梅承认他的这些想法还没有找到严格的证明,但是非常有信心地表示他一定会证出来.他甚至谦虚地表示,这些想法不能完全归功于自己,有些想法是他的同事刘维尔(Liouville)告诉他的.接下来,刘维尔则做了并不热情的发言.首先刘维尔说,在研究方程整数解问题中使用复数不是他的创造,欧拉和高斯等人早就这样做了.然后刘维尔认为,拉梅的证明思想中漏洞太大很难补上.接下来是法国另一个大数学家柯西(Cauchy)发言.柯西表示相信拉梅会成功,他还提醒与会者,早在 1846 年 10 月他本人就提交给法国科学院一份报告,指出证明费马猜想的一种思路,只是后来没有时间做下去.

法国科学院的这种会议每周都举行一次,在 3 月 8 日,15 日和 22 日的会议上,大家为此

事仍旧辩论不休.到了 26 日,刘维尔收到了德国数学家库默尔寄来的一封信,才终止了辩论.

库默尔在信中写到:拉梅的想法是需要环 $\mathbb{Z}[\zeta_p]$ 具有惟一因子分解特性才能严格证明的,而且证明并不如拉梅所想的那样简单.接下来库默尔在信中又说,早在 3 年前他就证明了当 $p=23$ 时,环 $\mathbb{Z}[\zeta_{23}]$ 并不具有惟一因子分解特性! 所以单纯用拉梅的思想不能完全证明费马猜想.库默尔在信中还附上他发表的文章.这一切使一群法国大数学家哑口无言.

恩斯特·库默尔(Ernst Kummer, 1810~1893)于 1810 年 1 月 29 日生于德国索拉乌(Sorau,现为波兰的扎雷).父亲是医生.在他童年时,拿破仑军队侵犯他的家乡,也带来了斑疹伤寒的流行.他的父亲死于此病,使库默尔的心灵受到很大创伤.他发誓尽力使他的祖国免遭法国人的侵犯,大学毕业后研究炮弹的弹道曲线问题,并且后来在柏林军事学院教弹道学并一直对军事有浓厚兴趣.父亲去世后,母亲在艰苦的生活中把他和他的哥哥抚养成人并给他们启蒙教育.1828 年他在哈勒大学专攻神学,后来受数学教授舍尔克(H. Scherk)的影响改学数学.1831 年获博士学位后,在中学教了 10 年的数学和物理.在沉重的中学教书之余,坚持数学研究.他的中学学生当中有后来的大数学家克

罗内克(Kronecker).1839年,在狄里赫利的推荐下他被选为柏林科学院通讯院士.1842年成为布雷斯劳大学数学教授,他的数学研究从函数论开始转向数论.1855年他接替了狄里赫利的柏林大学数学教授职位,并成为柏林科学院正式院士.1856年魏尔斯特拉斯(Weierstrass)到柏林大学任副教授,1861年他的学生克罗内克也以科学院院士身份在柏林大学任教,他们3人组织了德国重要的数学研究中心,组织讨论班,讲授几何学、力学、曲面论和数论,库默尔还在军事学院兼课.他对培养学生十分热心,数学家施瓦茨(Hermann Schwatz),康托尔(Cantor)和戈丹(Paul Gordan)都曾是他的博士生.后来他从事许多行政工作.1863至1878年他是柏林科学院物理数学部的秘书,1865~1866年为柏林大学哲学院院长,1868~1869年任柏林大学校长.1893年5月14日他患流感不治,平静地离开人世.

让我们介绍一下库默尔对于费马猜想的主要贡献.

(1) 库默尔严格地证明了:如果环$\mathbb{Z}[\zeta_p]$具有惟一因子分解性质,则费马猜想对$n=p$成立,即$x^p+y^p=z^p$没有正整数解.证明的起点仍是(7.2)式,并且(7.2)式右边的p个因子也确实是两两互素的.但是由惟一因子分解性

质并不能推出每个因子 $z-\zeta_p^i x$ 都是 $\mathbb{Z}[\zeta_i]$中某个数 α_i 的 p 次方 α_i^p，而只能推出与 α_i^p 相伴，即 $z-\zeta_p^i x=\alpha_i^p\varepsilon$，其中 ε 是 $\mathbb{Z}[\zeta_p]$中的可逆元素.

可逆元素因子的出现给问题增加了极大的复杂性，对于欧拉所研究的 $p=3$ 的情形，环 $\mathbb{Z}[\zeta_3]$只有 6 个可逆元素：± 1，$\pm\zeta_3$ 和 $\pm\zeta_3^2$. 但是当 $p\geqslant 5$ 时，$\mathbb{Z}[\zeta_p]$有无穷多个可逆元素！而且对比较大的素数 p，一直到今天人们还没有找出环 $\mathbb{Z}[\zeta_p]$的全部可逆元素，这是数论中一个重要的未解决问题. 然而库默尔发现了环 $\mathbb{Z}[\zeta_p]$中一些特殊的可逆元素 $\varepsilon_i=\dfrac{\zeta_p^i-1}{\zeta_p-1}(2\leqslant i\leqslant p-1)$，这些称为**分圆单位**（可逆元素也称为**单位**），"**分**圆"一词来源于：在复数平面中以原点为中心并且半径为 1 的单位圆上，p 个复数 $\zeta_p^0=1,\zeta_p,\zeta_p^2,\cdots,\zeta_p^{p-1}$ 恰好把圆周 n 等份.

要证明 $\varepsilon_i(2\leqslant i\leqslant p-1)$是可逆元素并不困难. 由于 $\varepsilon_i=\dfrac{\zeta_p^i-1}{\zeta_p-1}=\zeta_p^{i-1}+\zeta_p^{i-2}+\cdots+\zeta_p+1$，右边是用 1 和 ζ_p 进行加法和乘法运算得到的数，所以 $\varepsilon_i\in\mathbb{Z}[\zeta_p]$. 另一方面，当 $2\leqslant i\leqslant p-1$时 i 与 p 互素，所以存在整数 a 和 b，使得 $ia+pb=1$（定理 3.1）. 由 $\zeta_p^p=1$ 可知

$$\zeta_p = \zeta_p^1 = \zeta_p^{ia+pb} = \zeta_p^{ia} \cdot (\zeta_p^p)^b = \zeta_p^{ia}.$$

于是

$$\varepsilon_i^{-1} = \frac{\zeta_p - 1}{\zeta_p^i - 1} = \frac{\zeta_p^{ia} - 1}{\zeta_p^i - 1} = \zeta_p^{i(a-1)} + \zeta_p^{i(a-2)} + \cdots + \zeta_p^i + 1$$

也属于 $\mathbb{Z}[\zeta_p]$. 这表明 $\varepsilon_i(2 \leqslant i \leqslant p-1)$ 都是环 $\mathbb{Z}[\zeta_p]$ 中的可逆元素.

库默尔发现通过这些分圆单位,再利用乘除运算可以生成很多的可逆元素. 他详细研究了这些单位的性质,然后才能证明公式(7.2)不可能成立,从而最终证明了,若环 $\mathbb{Z}[\zeta_p]$ 有惟一因子分解性质,则费马猜想对 $n = p$ 时成立.

库默尔计算出:当 $p = 3, 5, 7, 11, 13, 17$ 和 19 时,环 $\mathbb{Z}[\zeta_p]$ 具有惟一因子分解特性,所以当 n 为这些素数时,费马猜想是成立的. 他用深刻和统一的方法得到的这个结果,是费马猜想的很大突破.

(2) 接下来,库默尔又发现环 $\mathbb{Z}[\zeta_{23}]$ 不具有惟一因子分解特性(他甚至猜想当 $p \geqslant 23$ 时, $\mathbb{Z}[\zeta_p]$ 均不具有惟一因子分解特性,这个猜想一直到 1976 年才被证明),所以进一步研究费马猜想又遇到困难. 库默尔并不死心,他把关于数的分解创造性地推广成所谓"理想数"的分

解.然后证明了环$\mathbb{Z}[\zeta_p]$中理想数的分解有惟一因子分解特性.利用理想数这个概念,库默尔研究了环$\mathbb{Z}[\zeta_p]$的另一个性质,称为环$\mathbb{Z}[\zeta_p]$的类数,表示成h_p.h_p是一个正整数,并且当$h_p=1$时,环$\mathbb{Z}[\zeta_p]$就有通常意义下的惟一因子分解特性,而当$h_p>1$时则不然.库默尔给出计算类数h_p的一种方法,用这种方法他算出当$p\leqslant 19$时,h_p均为1.而h_{23}等于3.所以环$\mathbb{Z}[\zeta_{23}]$没有惟一因子分解特性.库默尔计算h_p的办法中需要知道环$\mathbb{Z}[\zeta_p]$中全部可逆元素,所以对于大的素数p,计算类数h_p仍然非常困难.

接下来,库默尔证明了一个重要的结果.如果$h_p>1$,但是h_p不被p除尽,则费马猜想对于$n=p$的情形也仍然是对的.比如说:当$p=23$时$h_{23}=3$,所以环$\mathbb{Z}[\zeta_{23}]$没有惟一因子分解性质,但是3不被23除尽,所以费马猜想对于$n=23$的情形也成立,即$x^{23}+y^{23}=z^{23}$没有正整数解.

(3) 由于类数h_p很难计算,我们不能期望先算出h_p然后再看它是否被p除尽.库默尔随后又有新的发现,他找到了一种不用计算h_p的值就可以判断是否$p\mid h_p$的新方法,而且是非常初等的判别法.令$B_0=1$,然后用下面公式依

次算出一批有理数 $B_n(n=1,2,3,\cdots)$,

$$-(n+1)B_n=B_0+\binom{n+1}{1}B_1+\binom{n+1}{2}B_2+\cdots+\binom{n+1}{n-1}B_{n-1}.$$

例如取 $n=1$,则 $-2B_1=B_0=1$,所以 $B_1=-\frac{1}{2}$.又令 $n=2$,则 $-3B_2=B_0+\binom{3}{1}B_1=1+3\cdot\left(-\frac{1}{2}\right)=-\frac{1}{2}$,所以 $B_2=\frac{1}{6}$.注意符号$\binom{n}{m}$表示组合数$\frac{n!}{m!(n-m)!}$.继续下去

$$B_4=-\frac{1}{30},B_6=\frac{1}{42},B_8=-\frac{1}{30},$$

$$B_{10}=\frac{5}{66},B_{12}=-\frac{691}{2730},\cdots$$

而当 n 是大于 1 的奇数时,$B_n=0$,这些有理数称为**伯努利数**,因为在 17 世纪数学家伯努利(Jacob Bernoulli)就研究过这些数.库默尔对 $p\mid h_p$ 的初等判别法是说:

对每个奇素数 p,$p\nmid h_p$ 当且仅当 p 除不尽所有 $B_2,B_4,B_6,\cdots,B_{p-3}$(写成既约分数)的分子.

满足上面条件的奇素数称为**正规素数**,库默尔的结果(2)可以叙述成:当 p 是正规素数

时，费马猜想对 $n=p$ 成立．通过计算发现：在100以内的所有奇素数当中除了37、59和67之外，其余都是正规素数．这就证明了当 n 是这些奇素数时，费马猜想成立．对于 $p=37$，算出 B_{32} 的分子7709321041217可被37除尽，所以 $37|h_{37}$，即37是不正规素数．类似可算出59和67也是不正规素数．从正规素数和费马猜想的上述联系，人们自然希望正规素数愈多愈好．基于概率的考虑，猜想正规素数和不正规素数都有无穷多个，而且比例各占 $\sqrt{e}=61\%$ 和 $1-\sqrt{e}=39\%$（e 是自然对数的底数）．也就是说，猜想正规素数比不正规素数要多．但是目前只证明了不正规素数有无穷多个．至于正规素数是否有无穷多个，则至今没有答案．

除了上述三项结果与费马猜想有直接关系之外，库默尔还证明了另一些深刻的数论结果．比如说，他还给出计算类数 h_p 的一个公式，但是计算中需要环 $\mathbb{Z}[\zeta_p]$ 中可逆元素的知识．库默尔从费马猜想入手，深刻地研究了一类域的数论性质，这类域称为**分圆域**．而在库默尔之前半个世纪，高斯在研究方程 $ax^2+bxy+cy^2=n$（$x^2+y^2=n$）整数解时，深刻地探讨了二次域的数论性质．此后，他们的理论被德国数学家戴德金（Dedekind）和狄里赫利加以扩充和完善（比如说：库默尔的“理想数”变成为现在环论中

一个重要的概念:理想),形成了研究代数数域的一般理论.1898 年,德国大数学家希尔伯特写了《数论报告》一书,系统总结和进一步发展了代数数域的理论.从 19 世纪初期的高斯到 19 世纪末期的希尔伯特,数论的一个新的分支——代数数论走过了产生、发展和完善的历程,而费马猜想的研究对于这个新数论分支起了促进作用.另一方面,德国大数学家黎曼于 1859 年开创了数论的另一个重要分支:解析数论(我们在以后将作介绍),所以到了 19 世纪,数论的中心由法国转到德国.

到了世纪交替的 1900 年,希尔伯特在巴黎举行的第二次世界数学家大会上,提出了 23 个著名的数学问题.其中有 6 个数论问题,但是没有提到费马猜想.在 20 世纪,数论一直是一个十分活跃的研究领域.在数论中引入了非常丰富的数学思想、方法和工具,建立了许多新的数论分支,取得了一系列重大的数论成果.数学家和数学爱好者对费马猜想的研究也一直在进行.20 世纪 60 年代之后,计算机科学与技术得到飞速的发展,计算机在数学研究中得到愈来愈普遍的使用.1978 年,理论研究和计算机的使用相结合,人们证明了对于 n 是不超过 125 000的奇素数,费马猜想均成立.但是总起来说,20 世纪的前 80 年,费马猜想的研究没有

十分重大的突破.数学在不断的发展,这些发展事实上为费马猜想的最终解决在准备新的思想、方法和工具.但是在20世纪80年代以前人们并没有清楚地看到这一点.比如说,虽然证明了对所有不超过125 000的奇素数 p,方程 $x^p+y^p=z^p$ 都没有正整数解,但是人们仍旧不知道这样的奇素数 p 是否有无穷多个,这件事一直到80年代之后才被证明.所以在80年代以前,不少数学家认为费马猜想的最终解决是21世纪的事情.

到了20世纪80年代,对费马猜想的研究兴趣又重新高涨起来.如果说19世纪费马猜想的研究是代数和数论相互促进的结果,那么这一次则得益于几何与数论的结合.具体说来,1983年前联邦德国28岁的法廷斯(Faltings)证明了关于代数曲线上有理数点的莫代尔猜想,使人们感到几何方法或许是解决费马猜想的新的有效途径.这样一来,我们在下节要转到几何上来.

8 几何的介入：费马曲线

费马猜想是说：对每个 $n \geq 3$，方程 $x^n + y^n = z^n$ 没有正整数解．如果我们稍微扩大范围，在所有整数中考虑，那么此方程有一些平凡的整数解，即 x, y, z 当中至少有一个为 0 的那些解．当 n 为偶数时，这种解为 $(x, y, z) = (0, a, \pm a)$ 和 $(a, 0, \pm a)$．而当 n 为奇数时，这种解有 $(x, y, z) = (0, a, a)$，$(a, 0, a)$ 和 $(a, -a, 0)$，其中 a 是任意整数．除此之外，其他解都称为**非零**整数解（即 x, y, z 均为非零整数）．每个正整数解显然是非零整数解．反过来，由非零整数解也可造出正整数解．因为若 $(x, y, z) = (a, b, c)$ 是 $x^n + y^n = z^n$ 的非零整数解，则当 n 为偶数时，$(x, y, z) = (|a|, |b|, |c|)$ 就是正整数解．而当 n 是奇数时，$|a|$，$|b|$ 和 $|c|$ 一定彼此不同．取 z 为这三个绝对值当中的最大者，x 和 y 取为剩下两个绝对值，便得到方程的一组正整数解．所以费马猜想也可以说成：

对每个 $n \geq 3$，方程 $x^n + y^n = z^n$ 没有非零

整数解.

让我们再做一些变化.令 $X=\frac{x}{z}$, $Y=\frac{y}{z}$,则方程 $x^n+y^n=z^n$ 变成 $X^n+Y^n=1$.所以若 $(x,y,z)=(a,b,c)$ 是方程 $x^n+y^n=z^n$ 的非零整数解,则 $(X,Y)=\left(\frac{a}{c},\frac{b}{c}\right)$ 是方程 $X^n+Y^n=1$ 的非零有理数解.反之,若 $(X,Y)=(A,B)$ 是方程 $X^n+Y^n=1$ 的非零有理数解,将非零有理数 A 和 B 表成具有公分母的分数 $A=\frac{a}{c}$, $B=\frac{b}{c}$,其中 a,b,c 均为整数,则 $(x,y,z)=(a,b,c)$ 就是 $x^n+y^n=z^n$ 的非零整数解.于是,费马猜想又等价于说:

对每个 $n\geqslant 3$,方程 $X^n+Y^n=1$ 没有非零有理数解(指 X 和 Y 均不为 0).

我们把三个变量的费马方程变成只有两个变量的方程,并且解的范围也由整数环改成有理数域.一般来说,研究方程在域中的解比研究在环中的解要容易.如果把解的范围从有理数域再扩大到实数域,问题就变成几何学的研究对象了.因为对两个变量的实系数多项式 $f(x,y)$,方程 $f(x,y)=0$ 的全部实数解在坐标平面上通常绘成一条(或几条)连续曲线.比如前面的方程 $x^n+y^n-1=0$,当 n 为偶数时则是形如图 2 的封闭曲线,而当 n 为奇数时,是形如图 3

的曲线.

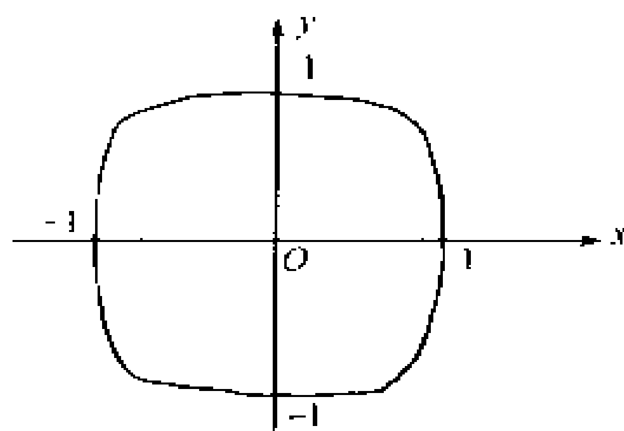


图 2　$x^4+y^4=1$

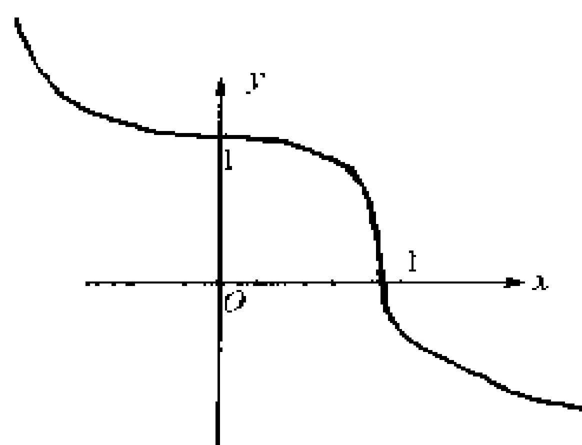


图 3　$x^3+y^3=1$

而费马猜想等价于说:这些曲线(称为**费马曲线**)上除了与坐标轴的交点之外,不再有其他有理数点.

对于一般情形,要从一条曲线上挑出所有的有理数点仍旧是相当困难的.但是对于特殊的曲线,我们可以用几何的方法做这件事.例如

对于单位圆周 $x^2+y^2=1$（图 4），我们取其上的一个有理点 $P=(-1,0)$. 若 (A,B) 是圆周上另一个有理点，将此点和点 P 连成直线 l，它的方程为 $y=t(x+1)$，其中 $t=\tan\alpha$ 为直线的斜率，α 是直线 l 与 x 轴正向的夹角（$-90°<\alpha<90°$），并且斜率 $t=\dfrac{B}{A+1}$ 是有理数. 反过来，对于过点 P 的任意直线 l，如果斜率 t 为有理数，该直线的方程为 $y=t(x+1)$. 此直线与单位圆周的交点就是方程组

$$\begin{cases} x^2+y^2=1, \\ y=t(x+1) \end{cases}$$

的解. 当 $t=0$ 时，直线 l 是 x 轴，它与单位圆周的交点为 $P=(-1,0)$ 和 $(1,0)$. 当 $t\neq 0$ 时令

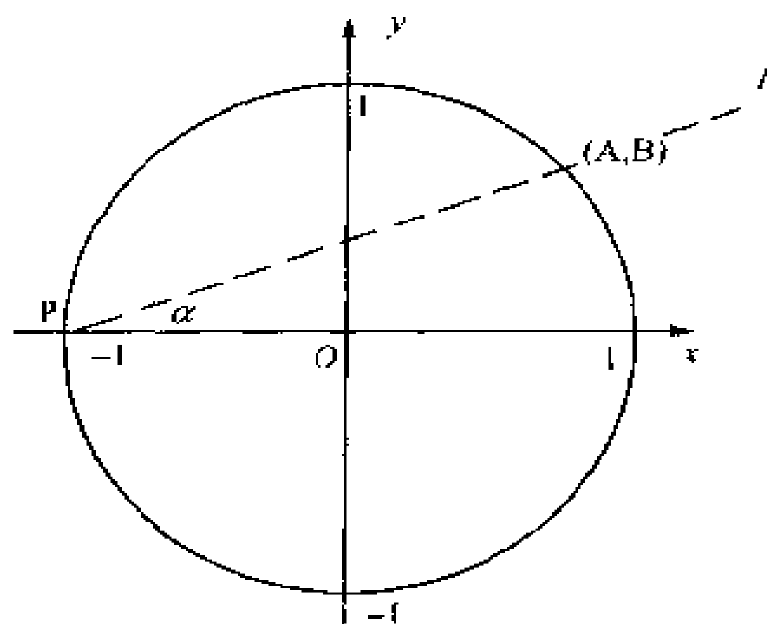


图 4

$s=t^{-1}$，则直线 l 可写成 $x+1=sy$. 将 $x=sy-1$ 代入 $x^2+y^2=1$ 得到 $(sy-1)^2+y^2=1$，即 $(s^2+1)y^2-2sy=0$. 于是有两个有理解 $y=0$ 和 $y=\frac{2s}{s^2+1}$. 当 $y=0$ 时 $x=sy-1=-1$，得到交点 P. 当 $y=\frac{2s}{s^2+1}$ 时 $x=sy-1=\frac{s^2-1}{s^2+1}$，便得到直线 l 与单位圆周的另一个交点 $(x,y)=\left(\frac{s^2-1}{s^2+1},\frac{2s}{s^2+1}\right)$. 取 s 为所有非零有理数，即直线 l 的斜率 $t=s^{-1}$ 跑遍所有非零有理数，我们便得到单位圆周上所有有理点（再加上 $(x,y)=(\pm1,0)$）. 所以方程 $x^2+y^2=1$ 的全部有理数解为 $(x,y)=\left(\frac{s^2-1}{s^2+1},\frac{2s}{s^2+1}\right)$（$s$ 过所有有理数）和 $(1,0)$. 将有理数 s 表成 $\frac{m}{n}$，其中 m 和 n 为整数，则

$$\frac{s^2-1}{s^2+1}=\frac{m^2-n^2}{m^2+n^2},\frac{2s}{s^2+1}=\frac{2mn}{m^2+n^2}.$$

可知 $(x,y,z)=(m^2-n^2,2mn,m^2+n^2)$ 是方程 $x^2+y^2=z^2$ 的整数解.

对于椭圆抛物线或双曲线，如果方程的系数都是有理数，那么知道其上的一个有理点，我们可以用同样的方法得到其上的无穷多个有理数点. 这些都是用几何方法寻求方程有理数解

的最简单例子.而对于平面上最简单的"曲"线:直线,容易判别其上是否有有理点.直线方程为 $ax+by=c$,如果其中 a,b,c 为有理数,并且 a 和 b 至少有一个不为零,易知它必有无穷多有理点.

综合上述,我们知道,设 $f(x,y)$ 是有理数为系数的多项式,当 $f(x,y)$ 的次数 $\leqslant 2$ 时,曲线 $f(x,y)=0$ 为直线或二次曲线(即椭圆,抛物线或双曲线).这些曲线上或者没有有理数点(例如由高斯定理知对每个正整数 n,方程 $x^2+y^2=3n^2$ 没有整数解,由此可知椭圆 $x^2+y^2=3$ 上没有有理数点),或者一定有无穷多个有理数点.这些曲线在代数几何学中归成一类,都称为**有理曲线**.用严格的几何学术语说,这类曲线是"亏格"为 0 的曲线.亏格这个概念来源于数学的一个分支:拓扑学.我们不打算讲述亏格的几何意义,只是告诉大家,每条曲线的亏格都是非负整数.

第二类曲线是亏格为 1 的曲线.这类曲线的典型方程为 $y^2=x^3+ax+b$,其中 a 和 b 均为整数,并且三次多项式 x^3+ax+b 在复数域中的三个根没有重根(熟知这相当于说:$4a^3+27b^2\neq 0$).这种曲线上可能有无穷多个有理点(例如 $y^2=x^3+17$ 就有 $(x,y)=(-2,3)$,$(5234,378661)$,$\left(\frac{137}{64},\frac{2651}{512}\right)$ 等无穷多有理

点)，也可能只有有限多个有理点(例如 $y^2=x^3+7x$ 只有一个有理点(0,0)，$y^2=x^3-9x$ 只有3个有理点$(x,y)=(0,0)$和$(\pm 3,0)$). 这类曲线称为**椭圆曲线**(注意不要把椭圆曲线和通常的椭圆混淆起来. 椭圆曲线这一名称的由来是：在历史上计算椭圆某段弧长时导致研究形如 $y^2=x^3+ax+b$ 的方程). 椭圆曲线上的所有有理点可以定义一种运算，使得这些点形成一个交换群. 20世纪以来，对于椭圆曲线上有理点群的研究不断深入，现在已成为内容丰富的一个数论分支，称为"椭圆曲线的算术理论". 我们以后将看到，费马猜想最终被证明是由于它与椭圆曲线的联系.

剩下的曲线是亏格≥2的曲线. 1923年，英国数学家莫德尔(Mordell)提出一个大胆的猜想：每条亏格≥2的曲线上都只有有限多个有理点. 我们说这个猜想是"大胆"的，因为当时没有多少例子来支持这个猜想. 对于亏格≥2的曲线，决定它的全部有理点是很困难的. 所以在此猜想提出之后，不少人持怀疑的态度. 另一方面，美国和苏联等地的几何学家们(特别是哈佛大学的代数几何学家和莫斯科大学沙弗列维奇领导的研究小组)为研究这个猜想作了不懈的努力. 在这些研究的基础上，1983年，德国28岁的法廷斯证明了莫德尔猜想是正确的，法廷

斯的这项工作于 *1986* 年得到世界数学的最高奖:菲尔兹(Fields)奖.

现在将法廷斯结果用于费马曲线 $x^n+y^n=1$.当 $n=1$ 和 2 时,$x+y=1$ 和 $x^2+y^2=1$ 分别是直线和单位圆周,所以均是亏格为 0 的有理曲线.事实上,曲线 $x^n+y^n=1$ 的亏格为 $\frac{1}{2}(n-1)(n-2)$.当 $n=1$ 和 2 时此曲线的亏格为 0(即有理曲线),我们已经知道这两条曲线上都有无穷多个有理点.当 $n=3$ 时,$x^3+y^3=1$的亏格为 1,即为椭圆曲线.为了具体地把 $x^3+y^3=1$ 变成椭圆曲线的典型形式,我们令 $y=\frac{v-9}{v}$, $x=\frac{3u}{v}\left(\text{也就是 } u=\frac{3x}{1-y}, v=\frac{9}{1-y}\right)$,则方程 $x^3+y^3=1$ 变成 $u^3=v^2-9v+27$.再令 $X=4u$, $Y=8v-36\left(\text{即 } u=\frac{X}{4}, v=\frac{Y+36}{8}\right)$,则方程就变成椭圆曲线的标准形式

$$Y^2=X^3-16\times 27.$$

欧拉已经证明了 $x^3+y^3=1$ 只有平凡的有理解 $(x,y)=(0,1)$和$(1,0)$,将此值代人 $X=4u=\frac{12x}{1-y}$ 和 $Y=8v-36=\frac{72}{1-y}-36=\frac{36(1+y)}{1-y}$,便知椭圆曲线 $Y^2=X^3-16\times 27$ 只有一组有理解$(X,Y)=(12,36)$.当 $n\geqslant 4$时,费马曲线

$x^n+y^n=1$ 的亏格 $\frac{1}{2}(n-1)(n-2)\geqslant 3$，于是由法廷斯的结果，可知 $x^n+y^n=1(n\geqslant 4)$ 只有有限多个有理解．这个结果还不能推出费马猜想，因为本节一开始就说过，费马猜想等价于说 $x^n+y^n=1$ **没有**非零的有理数解．

但是法廷斯结果还是给出费马猜想的一种推进．根据上述可知，$x^n+y^n=1(n\geqslant 4)$ 的非零有理数解只有有限多个(而我们的希望是：根本就没有非零有理数解)．对 $x^n+y^n=1$ 的每个非零整解 $x=\frac{a}{c}, y=\frac{b}{c}$，其中 a,b,c 是互素的非零整数，则 $(x,y,z)=(a,b,c)$ 就是 $x^n+y^n=z^n$ 的一组非零整数解．由 a,b,c 互素可知其中任意两个整数都是互素的(我们对 $n=2$ 的情况证明过这件事，$n\geqslant 3$ 情形可同样证明)．这样的整数解称为 $x^n+y^n=z^n$ 的**基本解**，而每个非零整数解都可由某个基本解共同乘上一个非零整数而得到．按照上述方式，可知 $x^n+y^n=1$ 的非零有理数解和 $x^n+y^n=z^n$ 的基本解是一一对应的．所以由法廷斯结果对费马猜想得出的结论为：当 $n\geqslant 4$ 时，方程 $x^n+y^n=z^n$ 的基本解 $(x,y,z)=(a,b,c)$(即 a,b,c 为两两互素的非零整数)至多有有限多个(而费马猜想则等价于说根本没有基本解)．

仅管法廷斯结果没有完全解决费马猜想，

而且也没有对任何新的 n 值证明费马猜想.但是法廷斯结果是对所有亏格 $\geqslant 2$ 曲线的论断,用到费马曲线 $x^n+y^n=1$ 上只不过是一个小小的推论.另一方面,这个推论是对所有的 $n\geqslant 4$ 都是适用的.在这之前,人们还没有对所有的 n 给出费马猜想的一般性结果.最重要的是:法廷斯结果对费马猜想的这个推动打破了百余年来费马猜想研究工作的沉寂局面.许多数学家相信,采用进一步的几何方法有希望解决费马猜想.短短的几年里,人们设计出三种以上的几何学研究方案.但是得到成功的却是在这些方案之外,对费马猜想和椭圆曲线相联系的一个意外的发现.而且这种联系是通过研究数论的另一个重要手段——解析方法建立起来的.所以我们在下一节先介绍研究数论的解析方法,然后再介绍椭圆曲线的知识和与费马猜想的意外联系.

9 解析的介入

整数等间隔地排列在实数轴上，所以数论本质上是研究离散对象性质的一门数学学问．17 世纪由牛顿和莱布尼茨发明了微积分，后来又发展出复变函数论、调和分析等各种数学解析理论，都是以连续的对象作为研究主体．把解析的方法引到数论中来，是件非常有趣的事件，体现出世上离散对象和连续对象的联系．

第一个做这件事的是欧拉．18 世纪初期，他所热心的一个研究课题是把无穷多个实数 a_1，a_2，a_3，…，a_n，…全部加起来是否有极限．也就是说，令 $s_n = a_1 + a_2 + \cdots + a_n$，我们把实数序列前 n 个数相加 s_n 也表示成 $\sum_{i=1}^{n} a_i$．要问当 n 趋于无穷时(表示成 $n \to \infty$)，实数 s_n 是否存在极限．如果有极限为 A，即 $\lim_{n \to \infty} s_n = A$，我们就把这个极限记成 $\sum_{i=1}^{\infty} a_n$．并且说这个求和是**收敛的**(收敛于 A)．这类数学问题称为**级数的求和**．

中学里学过的等比级数 $a_0=a, a_1=aq, a_2=aq^2, \cdots, a_n=aq^n, \cdots$，就是一个例子（其中 a 和 q 均为实数，称为等比级数的首项和公比）. 我们知道，对于这种情形当 $q\neq 1$ 时，

$$s_n = a_0 + a_1 + \cdots + a_n = \sum_{i=0}^{n} a_i$$

$$= a(1+q+q^2+\cdots+q^n) = \frac{a(1-q^{n+1})}{1-q}.$$

如果 $|q|<1$，则当 $n\to\infty$ 时 $q^n\to 0$，于是 $s_n\to \frac{a}{1-q}$. 也就是说，

$$\sum_{i=0}^{\infty} a_i = \frac{a}{1-q},$$

当 $|q|>1$ 时，级数不收敛（也称为发散）. 另一个发散的典型例子为

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\cdots$$

将第 3，4 项相加则 $\frac{1}{3}+\frac{1}{4}>\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$，其后的 4 项相加，则 $\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}>\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{1}{2}$. 继续下去，再将其后的 8 项相加，又得到

$\frac{1}{9}+\frac{1}{10}+\cdots+\frac{1}{16}>8\cdot\frac{1}{16}=\frac{1}{2}$.由此可知这个级数是无穷多个$\frac{1}{2}$加在一起,所以是无穷大$\infty$,即$\sum_{n=1}^{\infty}\frac{1}{n}=\infty$(不收敛).接下来欧拉考虑$\sum_{n=1}^{\infty}\frac{1}{n^s}=\sum_{n=1}^{\infty}n^{-s}$.欧拉发现,当$s$比1稍微大一点时,则这个级数就收敛.这个求和值显然依赖于s,所以我们得到一个关于s的实函数

$$\zeta(s)=\sum_{n=1}^{\infty}n^{-s}\quad(s>1).$$

类似地,欧拉还考虑实数序列$a_1,a_2,\cdots,a_n,\cdots$乘积是否有极限.以$\prod_{i=1}^{n}a_i$表示序列前$n$个实数之乘积$a_1a_2\cdots a_n$,要问极限$\lim_{n\to\infty}\prod_{i=1}^{n}a_i$是否存在.如果这个极限存在并且是$A$,我们记为$\prod_{i=1}^{\infty}a_i=A$,并且当$A\neq 0$时,称无穷乘积是收敛的.

有限个实数相加或相乘是满足交换律的,即计算$a_1+a_2+\cdots+a_n$或$a_1a_2\cdots a_n$时可以任意交换$a_1,\cdots,a_n$的次序.无穷级数求和$a_1+a_2+\cdots+a_n+\cdots$与无穷乘积$\prod_{i=1}^{\infty}a_i$在调换因子

次序时要格外小心.例如可以算出

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots+\frac{1}{2n-1}$$

$$-\frac{1}{2n}+\cdots=\ln 2, \tag{9.1}$$

其中右边的自然对数 ln2 是正实数.但是如果把左边级数各项出现的次序改成

$$1-\frac{1}{2}-\frac{1}{2^2}-\cdots-\frac{1}{2^n}\cdots+\frac{1}{3}-\frac{1}{2\cdot 3}$$

$$-\frac{1}{2^2\cdot 3}-\cdots-\frac{1}{2^n\cdot 3}-\cdots+\frac{1}{5}-\frac{1}{2\cdot 5}$$

$$-\frac{1}{2^2\cdot 5}-\cdots-\frac{1}{2^n\cdot 5}-\cdots+\cdots \tag{9.2}$$

在(9.1)式的左边,所有正奇数的倒数 $1,\frac{1}{3},\frac{1}{5}$,…都是正号,而(9.2)式中每行的第 1 个数正好是它们.在(9.1)式的左边,所有正偶数的倒数 $\frac{1}{2},\frac{1}{4},\frac{1}{6}$,…都是负号,由于每个偶数都惟一地写成正奇数乘以 2^l($l\geqslant 1$),可知(9.2)式带负号的那些项也正好与(9.1)式左边的诸项一致.但是不难算出,(9.2)式的每一行的值都是 0,所以(9.2)式的值为 0,而 $\ln 2>0$.这表明若 a_1+a_2

$+\cdots+a_n+\cdots$收敛于 A,即 $\sum_{n=1}^{\infty} a_n = A$,将诸 a_n 重新排列次序之后,其和可能不存在极限,或者极限不等于 A.

进一步的研究表明,如果对 $a_1+a_2+\cdots+a_n+\cdots$加上一些条件,就可以克服这些麻烦.通常用得最多的条件是要求 $\sum_{i=1}^{\infty} |a_i| = |a_1| + |a_2| + \cdots + |a_n| + \cdots$是收敛的,这称 $\sum_{i=1}^{\infty} a_i$ 是**绝对收敛**.可以证明,在绝对收敛条件下,若 $\sum_{i=1}^{\infty} a_i = A$,则将 a_i 任意交换次序,其和仍有极限,并且极限值仍为 A.上面例子出现的问题,是由于 $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots$不绝对收敛,即 $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots=\infty$(不收敛).

类似地,将分配律$(a_1+a_2)(b_1+b_2)=a_1b_1+a_2b_1+a_1b_2+a_2b_2$ 试图推广到 $(\sum_{i=1}^{\infty} a_i)(\sum_{j=1}^{\infty} b_j)$ 上去也会有同样的麻烦.假如 $\sum_{i=1}^{\infty} a_i$ 和 $\sum_{j=1}^{\infty} b_j$ 都是收敛的,其和值分别为 A 和 B.我们如何对

$$\left(\sum_{i=1}^{\infty} a_i\right)\left(\sum_{j=1}^{\infty} b_j\right) = (a_1 + a_2 + \cdots$$

$$+ a_n + \cdots)(b_1 + b_2 + \cdots + b_m + \cdots) \quad (9.3)$$

采用分配律？每个 a_n 和每个 b_m 相乘得到所有的 $a_n b_m (n, m = 1,2,3,\cdots)$，我们甚至都很难说出这无穷多项以何种排列方式相加，才是最合理的. 在这里，又是绝对收敛条件可以解决麻烦：如果 $\sum_{i=1}^{\infty} |a_i|$ 和 $\sum_{j=1}^{\infty} |b_j|$ 均收敛，那么可以证明，无论将无穷多项 $a_n b_m (n, m = 1,2,3,\cdots)$ 如何排列，其和一定有极限，并且极限值一定为(9.3)式的左边 $\left(\sum_{i=1}^{\infty} a_i\right)\left(\sum_{j=1}^{\infty} b_j\right) = AB$.

类似地，如果 3 个，4 个，以至任意有限多个无限求和都是绝对收敛的，那么也可以使用分配律来求它们的乘积. 最后，对于无限多个无限求和是否有分配律？我们以后只用到以下的特殊情形：

$$(1 + a_1 + a_2 + \cdots)(1 + b_1 + b_2 + \cdots)$$
$$\cdot (1 + c_1 + c_2 + \cdots)\cdots(\text{无穷个}) = ? \quad (9.4)$$

假设左边每个括号均是收敛的，令 $1 + a_1 + a_2$

$+\cdots=A$, $1+b_1+b_2+\cdots=B$, $1+c_1+c_2+\cdots=C$,如此等等,并且无穷乘积 $ABC\cdots$ 也是收敛的.我们如何对(9.4)式左边使用分配律?首先我们要弄清(9.4)式左边展开之后都有哪些项.我们知道,(9.4)式的左边是

$$A=1+a_1+a_2+\cdots,$$

$$AB=(1+a_1+a_2+\cdots)(1+b_1+b_2+\cdots)$$

$$=1+a_1+a_2+\cdots+b_1+b_2+\cdots+\sum_{i,j}a_ib_j,$$

$$ABC=(1+a_1+a_2+\cdots)$$

$$\cdot(1+b_1+b_2+\cdots)(1+c_1+c_2+\cdots)$$

$$=1+\sum a_i+\sum b_j+\sum c_k+\sum a_ib_j+\sum a_ic_k$$

$$+\sum b_jc_k+\sum a_ib_jc_k$$

....

的极限.若$(1+a_1+a_2+\cdots)$,$(1+b_1+b_2+\cdots)$,$(1+c_1+c_2+\cdots)$等等都是绝对收敛的,由前所述,上述关于 A,AB,$ABC\cdots$的诸式均正确,所以它们的极限就是无穷乘积 $ABC\cdots$.

初看起来,将(9.4)式左边使用分配律,似乎应当从每个括号中取**任意**一项,然后将它们乘起来,再把无穷多个这种乘积加在一起,这样做就会有许多乘积是无穷乘积 $a_ib_jc_k\cdots$.但事

实上并不如此，因为在 $A, AB, ABC, \cdots$ 的展开式中，每项都是**有限乘积**：$1, a_i, b_j, c_k, \cdots, a_ib_j, a_ic_k, b_jc_k, \cdots, a_ib_jc_k, \cdots$. 取它们的极限，可知(9.4)式左边按分配律展开应当是只在**有限**多个括号中取不是1的数，而其余括号均取求和的第1项1，这样乘起来都是有限乘积，把这些有限乘积(以任意次序)相加就是 $ABC\cdots$. 这是我们要特别说明的一点.

有了以上这些准备，就可继续讲述我们的故事. 1737年，欧拉在研究无限求和以及无限乘积的许多新奇的现象时，发现了如下的一个等式：

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} + \cdots$$

$$= \prod_p (1 - p^{-s})^{-1} \quad (s > 1), \tag{9.5}$$

其中左边是无穷级数求和，而右边 p 过所有的素数，所以是无穷乘积. 由于(9.5)式诸项都是正数，从而当 $s > 1$ 时绝对收敛性没有问题，所以我们可以采用上述的分配律并且求和可采取任意次序. 现在解释公式(9.5)为什么是正确的. 对每个素数 p，

$$(1 - p^{-s})^{-1} = \frac{1}{1 - p^{-s}} = 1 + p^{-s} + p^{-2s} + \cdots$$

$$+ p^{-ns} + \cdots$$

注意此式右边的等比级数是(绝对)收敛的.于是(9.6)式可写成

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} + \cdots$$

$$= \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{ns}} + \cdots\right)(s > 1), \tag{9.6}$$

此式右边就是无限求和的无穷乘积.按照前面的说明,将(9.6)式右边按分配律展开时,应当只在有限个括号中取不为 1 的项,其余括号均取 1.假设在素数 p_1 对应的括号中取 $\frac{1}{p^{n_1 s}}$, p_2 对应的括号中取 $\frac{1}{p_2^{n_2 s}}$, $\cdots$, p_t 对应的括号中取 $\frac{1}{p_t^{n_t s}}$,而其余括号均取第 1 项 1,乘起来为

$$\frac{1}{p_1^{n_1 s} p_2^{n_2 s} \cdots p_t^{n_t s}} = \frac{1}{n^s},$$

其中 $n = p_1^{n_1} p_2^{n_2} \cdots p_t^{n_t}$. $\frac{1}{n^s}$ 是(9.6)式左边的一项,由于正整数的惟一因子分解性,(9.6)式左边求和的各项与右边展开后的各项是一一对应的.当 $s > 1$ 时,绝对收敛性保证求和与各项排

列次序无关．这就证明(9.6)式(和(9.5)式)的正确性．

公式(9.5)和(9.6)称为函数 $\zeta(s)$ 的欧拉无穷乘积表达式，它体现了整数环 $\mathbb{Z}$ 的唯一因子分解特性．我们还可用公式(9.5)证明数论的其他性质．举一个最平凡的例子，就是用来证明素数有无穷多个：如果素数只有有限多个，则(9.5)式的右边是有限个因子 $(1-p^{-s})^{-1}$ 的乘积．取 $s=1$，右边是有限个正数相乘，所以值为正有理数．但是左边当 $s=1$ 时其值为 ∞．这就导出矛盾．这个推导虽然简单，而且证明的结果也是2000年前就知道的事实，但沿着这种思路可以得出不平凡的结果．在欧拉之后，狄里赫利考虑满足如下性质的函数($\mathbb{C}$ 表示复数全体，m 是固定的正整数)

$$\chi:\mathbb{Z}\to\mathbb{C}.$$

(1) 如果 $a\equiv b(\mathrm{mod}\,m)$，则 $\chi(a)=\chi(b)$．

(2) 当 a 与 m 互素时，$\chi(a)=0$．否则 $\chi(a)\neq 0$．

(3) 对任意整数 a 和 b，$\chi(ab)=\chi(a)\cdot\chi(b)$．

可以证明这样的函数一共有 $\varphi(m)$ 个($\varphi(m)$ 是欧拉函数)．例如，

$$\chi_0(a) = \begin{cases} 0, 如果 (a, m) > 1, \\ 1, 如果 (a, m) = 1 \end{cases}$$

就是这样的函数.对于 $m = 4, \varphi(4) = 2$.除了 χ_0 之外,模 4 的另一个这样的函数为

$$\chi(a) = \begin{cases} 0, & 若 (a,4) > 1 (即 2 \mid a), \\ 1, & 若 a \equiv 1 (\bmod 4), \\ -1, & 若 a \equiv -1 (\bmod 4). \end{cases}$$

对于每个这样的函数 χ(称为模 m 的**狄里赫利特征**),可以构作级数(由性质(3)可知 $\chi(1) = 1$)

$$L(s, \chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s}$$

$$= 1 + \frac{\chi(2)}{2^s} + \frac{\chi(3)}{3^s} + \cdots + \frac{\chi(n)}{n^s} + \cdots \tag{9.7}$$

由性质(3)可知这个函数也有欧拉乘积表达式:

$$L(s, \chi) = \prod_p \left(1 + \frac{\chi(p)}{p^s} + \frac{\chi(p)^2}{p^{2s}} + \cdots + \frac{\chi(p)^n}{p^{ns}} + \cdots\right)$$

$$= \prod_p (1 - \chi(p) p^{-s})^{-1} \quad (s > 1). \tag{9.8}$$

$L(s,\chi)$称为狄里赫利 L 函数. 可以证明(9.7)式在$s > 1$时是绝对收敛的. 所以(9.8)式对于$s > 1$时成立. 狄里赫利还证明了对每个 $\chi \neq \chi_0$, $L(s,\chi)$在 $s = 1$ 时也收敛,且其值 $L(1,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n}$不为 0. 利用 L 函数的这个性质,狄里赫利证明了如下的深刻结果:对每个与 m 互素的正整数 a,等差级数,$a, a+m, a+2m, \cdots, a+lm, a+(l+1)m, \cdots$当中一定有无穷多个素数. 这不仅在当时是一个重要的新结果,而且现在用纯数论方法来证明也是很困难的.

欧拉只是对实数值 s 来研究实函数 $\zeta(s)$. 1859 年,德国大数学家黎曼(Riemann, 1826 ~ 1866)做了重要的工作,他把 $\zeta(s)$看成是复变量 $s = \sigma + it$ 的函数. 级数 $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ 在 s 的实数部分 $\mathrm{Re}(s) = \sigma$ 大于 1 时是绝对收敛的,因为$|n^{-s}| = n^{-\sigma}$. 并且把 $\zeta(s)$解析开拓成整个复平面上的亚纯函数,而且还证明了$\zeta(s)$满足一个函数方程

$$\zeta(s) = f(s)\zeta(1-s),$$

其中 $f(s)$ 是一个熟知的复变函数. 在把 $\zeta(s)$ 的变量 s 扩大到整个复平面上之后, 证明了 $s=-2,-4,-6,\cdots$ 都是 $\zeta(s)$ 的零点 (即 $\zeta(-2n)=0, n=1,2,3,\cdots$). 这些称为 $\zeta(s)$ 的**平凡零点**, 黎曼猜想: $\zeta(s)$ 的所有其他零点的实数部分都等于 $\frac{1}{2}$. 这就是著名的黎曼猜想. 黎曼猜想可以推出许多数论上的重要结果, 但是黎曼猜想至今未被证明(或推翻).

基于黎曼这项工作, 开创了用复变函数的解析方法进行数论研究的途径. 产生了数论的一个重要分支: 解析数论. 函数 $\zeta(s)$ 也由此被后人称作**黎曼 ζ(*zeta*) 函数**. 类似地, 人们发现狄里赫利 L 函数 $L(s,\chi)$ 也可以解析开拓成整个复平面上的函数, 并且也有形如 $L(s,\chi)=f(s)L(1-s,\bar{\chi})$ 的函数方程, 其中 $\bar{\chi}$ 是与 χ 有关的另一个狄里赫利特征.

为了本书后面的需要, 我们再介绍复变函数的一个解析特性, 并且说明这种解析特性和数论的联系. 让我们从大家熟悉的多项式情形谈起. 设 $f(z)$ 是一个复系数的 n 次多项式, 代数基本定理是说, $f(z)$ 在复数域中一共有 n 个根 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 并且 $f(z)$ 可以写成

$$f(z)=a(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_n),$$

其中 $a\neq 0$ 是多项式 $f(z)$ 的最高次项的系数

(首项系数).这 n 个根当中可能有相同的重根.假设 n 个根当中恰有 t 个为 $\alpha(t\geqslant 1)$,称 α 是 $f(z)$ 的 n 重根或 **n 重零点**这时

$$f(z) = a(z-\alpha)^t g(z),$$

其中 $g(z)$ 不再有根 α,因此 $g(\alpha)\neq 0$.所以 t 可由 $\lim\limits_{z\to\alpha}\frac{f(z)}{(z-\alpha)^t} = ag(\alpha)\neq 0$ 来决定.现在设 $f(z)$ 是有理函数

$$f(z) = \frac{A(z)}{B(z)},$$

其中 $A(z)$ 和 $B(z)$ 是复系数的多项式,并且 $A(z)$ 和 $B(z)$ 没有公共根.则 $B(z)$ 的 t 阶零点 $\beta(t\geqslant 1)$ 称为 $f(z)$ 的 **t 阶极点**.这时 $B(z) = (z-\beta)^t g(z)$,$g(\beta)\neq 0$.由于 $A(z)$ 和 $B(z)$ 没有公根,$A(\beta)\neq 0$.于是 t 由

$$\lim_{z\to\beta}(z-\beta)^t f(z) = \frac{A(\beta)}{g(\beta)}\neq 0$$

所决定.通过这些直观,我们可以考虑任何一个复变函数 $f(z)$.对某个复数 α,如果有正整数 t,使得 $\lim\limits_{z\to\alpha}\frac{f(z)}{(z-\alpha)^t}$ 是非零复数,称 α 是 $f(z)$ 的 **t 阶零点**.如果 $\lim\limits_{z\to\alpha}(z-\alpha)^t f(z)$ 是非零复数,称 t 是 $f(z)$ 的 **t 阶极点**.

零点和极点是复变函数的解析性质，这种性质也可用于研究数论问题．例如对黎曼 ζ 函数 $\zeta(s)$，可以算出 $\lim_{s\to 1}(s-1)\zeta(s)=1$，所以 $s=1$ 是 $\zeta(s)$ 的 1 阶极点．并且由函数方程可推出 $\zeta(s)$ 不再有任何极点．由此可以得到素数分布的一些重要结果，解析方法也可用到代数数论的研究中．对于每个代数数域 K，狄德金在（19 世纪）构作了一个 ζ 函数 $\zeta_K(s)$．它也有欧拉乘积展开（当 $\sigma>1$ 时），可解析开拓到整个复平面上，并且有形如 $\zeta_K(s)=f(s)\zeta_K(1-s)$ 的函数方程．在 $s=1$ 处人们计算出

$$\lim_{s\to 1}(s-1)\zeta_K(s)=\frac{2^{r_1}(2\pi)^{r_2}h_K R_K}{\sqrt{|d_K|}}\neq 0, \tag{9.9}$$

其中 r_1, r_2, h_K, R_K, d_K 都是与数域 K 的算术和代数性质有关的复数，比如 h_K 是数域 K 的类数．它是一个正整数，并且 $h_K=1$ 当且仅当数域 K 的整数环具有惟一因子分解特性．所以 (9.9) 式不仅表明 $s=1$ 是 $\zeta_K(s)$ 的 1 阶极点，而且计算左边的极限可以用来求出类数 h_K 的值．这就使得判别一个环是否具有惟一因子分解特性这个数论问题，归结于函数 $\zeta_K(s)$ 的一个极限值 $\lim_{s\to 1}(s-1)\zeta_K(s)$ 的计算问题．高斯和

库默尔就是用这种方法分别给出了二次域和分圆域的类数解析计算公式，用此来判别环 $\mathbb{Z}[\sqrt{d}]$和 $\mathbb{Z}[\zeta_m]$是否有惟一因子分解特性.

黎曼 ζ 函数

$$\zeta(s)=\sum_{n=1}^{\infty} n^{-s} \qquad (s=\sigma+it, \sigma>1) \tag{9.10}$$

在 $\sigma>1$ 时可以用计算右边的级数来求出 $\zeta(s)$ 的值.特别对正整数 $s=2,3,4,\cdots$可以这样做.但是这个级数的确切值求起来并不容易.欧拉于 1740 年用三角函数的性质，计算出 $\zeta(s)$在正偶数 $s=2,4,6,\cdots$处的值为

$$\zeta(2k)=\frac{(-1)^{k+1}(2\pi)^{2k}}{2(2k)!}B_{2k} \quad (k=1,2,3,\cdots), \tag{9.11}$$

其中 B_{2k}是第 7 节所述的伯努利数.例如当 $k=1,2$ 时，$B_2=\frac{1}{6}$，$B_4=-\frac{1}{30}$.于是

$$\begin{aligned}\zeta(2)&=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\frac{1}{36}+\frac{1}{49}+\cdots\\&=\frac{(2\pi)^2}{2\cdot(2!)}\cdot\frac{1}{6}=\frac{\pi^2}{6},\end{aligned}$$

$$\zeta(3) = 1 + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{4^3} + \cdots$$

$$= \frac{(2\pi)^4}{2 \cdot (4!)} \cdot \frac{1}{30} = \frac{\pi^4}{90},$$

等等.所以 $\zeta(2k)$是 π^{2k}乘上一个有理数(注意伯努利数都是有理数).圆周率 π 不仅不是有理数,而且也不是一般的无理数.π 不是任何整系数多项式的根.像$\sqrt{2}$,$\zeta_m = e^{\frac{2\pi i}{m}}$这样的无理数分别为整系数多项式 $x^2 - 2$ 和 $x^m - 1$ 的根,这种数称为**代数数**.其他无理数称为**超越数**.π 和 π 的任何方幂 π^l 都是超越数.所以从我们算出的公式(9.11)可知 $\zeta(2k)(k = 1,2,3,\cdots)$都是超越数.另一方面,$\zeta(s)$在正奇数 $s = 3,5,7,\cdots$处的值到目前为止没有像正偶数情形那样令人愉快的公式(9.11).人们猜想 $\zeta(3)$,$\zeta(5)$,$\zeta(7)$,…也都是超越数.这个问题至今没有解决.目前只知道 $\zeta(3) = 1 + \frac{1}{8} + \frac{1}{27} + \frac{1}{64} + \frac{1}{125} + \cdots$是无理数.

关于黎曼 ζ 函数 $\zeta(s)$在负整数处的值 $\zeta(-k)(k = 1,2,\cdots)$,我们不能用级数(9.10)来计算,因为在负整数处的值不是由(9.10)式定义的,而是通过函数方程把 $\zeta(s)$开拓到这些点上去.利用函数方程可算出

$$\zeta(-k) = -\frac{B_{k-1}}{k+1} \quad (k = 1,2,3,\cdots),$$

$$(9.12)$$

当 n 为大于 1 的奇数时，$B_n = 0$. 由此可知 $s = -2, -4, -6, \cdots$ 都是 $\zeta(s)$ 的零点. 而当 $s = -1, -3, -5, \cdots$ 时 $\zeta(s)$ 是非零的有理数. 注意 $\zeta(s)$ 在整数 $s \in \mathbb{Z}$ 的取值公式(9.11)和(9.12)中都出现了伯努利数 B_n，我们在第 7 节知道这种数具有数论意义，因为库默尔用它们来判别分圆域 $\mathbb{Q}(\zeta_p)$ 的类数 h_p 是否被 p 除尽，然后用于研究费马猜想. 也就是说，$\zeta(s)$ 在某些特殊点的取值有数论意义. 这也体现了解析函数和数论的联系.

从此以后，数论研究中就出现了一个十分吸引人的课题：对于某个数论对象 S（S 可以是整数环，$\mathbb{Z}[\sqrt{d}]$，$\mathbb{Z}[\zeta_m]$，也可以是费马曲线，椭圆曲线等等），是否可以构作出一个与 S 有关的复变函数 $f_S(z)$（通常称为 S 的 ζ 函数或 L 函数），使得 $f_S(z)$ 的各种解析性质（零点和极点特性，无穷乘积展开，函数方程，函数在特殊点的取值等等），能够反映对象 S 的数论性质. 这种研究在近 50 年里成为现代数论的一个热门课题. 我们在随后三节里再介绍这方面的两个例子. 一个是在平方和问题中引入复变

函数，称为**模形式**.另一个是在研究椭圆曲线 E 时引入一种复变函数 $L_E(s)$，称为椭圆曲线 E 的 L 函数.而这两类函数之间的联系最终导致费马猜想的完全解决.

10 平方和与模形式

让我们再回过头来谈平方和问题.我们在第6节介绍了费马对二平方和的猜想、欧拉对这猜想的证明以及高斯对二平方和问题的完全解决(定理6.4)、利用高斯整数环的惟一因子分解特性,高斯还得到方程 $x^2+y^2=n$ 的整数解个数的公式(定理6.9).我们把这个公式再写一遍:以 $N_2(n)$ 表示方程 $x^2+y^2=n$ 的整数解个数.令

$$n = 2^{a_0} p_1^{a_1} \cdots p_s^{a_s} q_1^{b_1} \cdots q_t^{b_t}, \quad (10.1)$$

其中 $p_1, \cdots, p_s, q_1, \cdots, q_t$ 是不同的奇素数,并且

$$p_i \equiv 1 \pmod 4 \quad (1 \leqslant i \leqslant s),$$

$$q_j \equiv 3 \pmod 4 \quad (1 \leqslant j \leqslant t),$$

则当 $b_1, \cdots, b_t$ 当中至少有一个为奇数时,$N_2(n)=0$(即 $x^2+y^2=n$ 无整数解),而当 $b_1, \cdots, b_t$ 均为偶数时

$$N_2(n) = 4(a_1+1)(a_2+1)\cdots(a_s+1). \tag{10.2}$$

上述结论可以变换一个说法. 我们以 $d_1(n)$表示 n 的被 4 除余 1 的正奇因子的个数, $d_3(n)$表示 n 的被 4 除余 3 的正奇因子的个数. 我们要证明

引理 10.1 对每个正整数 n, 方程 $x^2+y^2=n$ 的整数解个数为 $N_2(n)=4(d_1(n)-d_3(n))$.

证 设 n 的分解式为(10.1). 如果某个 b_i 是奇数, 不妨设 b_1 为奇数. 则 n 的每个奇因子都可写成 Aq_1^c, 其中 A 是与 q_1 互素的奇数, 而 $c=0,1,2,\cdots,b_1$. 由于 $q_1\equiv 3 \pmod 4$, 所以对每个固定的 A, 在 b_1+1 个奇因子 $Aq_1^c (0\leqslant c\leqslant b_1)$当中有一半模 4 同余于 1, 而另一半模 4 同余于 3. 所以总起来, n 的正奇因子中模 4 同余于 1 的个数 $d_1(n)$等于模 4 同余于 3 的个数 $d_3(n)$. 即 $4(d_1(n)-d_3(n))=0$, 而 $N_2(n)$也等于 0. 所以结论成立.

现在设 $b_1,\cdots,b_t$ 都是偶数. 则 n 的每个正奇因子有形式

$$d = p_1^{x_1}\cdots p_s^{x_s} q_1^{y_1}\cdots q_t^{y_t}, \tag{10.3}$$

其中 $0\leqslant x_i\leqslant a_i (1\leqslant i\leqslant s)$, $0\leqslant y_j\leqslant b_j (1\leqslant j\leqslant$

t). 由于 $p_i \equiv 1 \pmod 4$, $q_j \equiv 3 \equiv -1 \pmod 4$, 所以

$$d \equiv (-1)^{y_1+\cdots+y_t} \pmod 4. \quad (10.4)$$

每个 x_i 有 a_i+1 个取值, 所以 $p_1^{x_1}p_2^{x_2}\cdots p_s^{x_s}$ 共有 $(a_1+1)\cdots(a_s+1)$ 个可能. 对于固定的 $p_1^{x_1}\cdots p_s^{x_s}$, 则由(10.4)式知 $d \equiv 1 \pmod 4$ 当且仅当 $(-1)^{y_1+\cdots+y_t}=1$, 于是 $d \equiv 3 \pmod 4$ 当且仅当 $(-1)^{y_1+\cdots+y_t}=-1$. 所以对每个固定的 $p_1^{x_1}\cdots p_s^{x_s}$, 则形如(10.3)式的奇因子 d 当中模 4 余 1 的个数减去模 4 余 3 的个数应当为

$$\sum_{y_1=0}^{b_1}\cdots\sum_{y_t=0}^{b_t}(-1)^{y_1+\cdots+y_t}$$

$$=\left(\sum_{y_1=0}^{b_1}(-1)^{y_1}\right)\cdots\left(\sum_{y_t=0}^{b_t}(-1)^{y_t}\right).$$

由于 b_j 都是偶数, 可知上式每个括号的值均为 1. 这就表示对每个固定的 $p_1^{x_1}\cdots p_s^{x_s}$, 形如(10.3)式的奇因子 d 当中模 4 余 1 的比模 4 余 3 的均多 1 个. 而 $p_1^{x_1}\cdots p_s^{x_s}$ 共有 $(a_1+1)\cdots(a_s+1)$ 个. 这就表明 $4(d_1(n)-d_3(n))=4(a_1+1)\cdots(a_s+1)$, 而 $N_2(n)$ 也等于此数, 证毕.

很自然地, 接下来要考虑一个正整数 n 是否可以表成三平方和, 即方程 $x^2+y^2+z^2=n$

是否有整数解.考虑模 8 同余,可知 7 不能表成三平方和,哪些正整数可以表成三平方和这个问题也是由高斯解决的.他的结论是:

引理 10.2(高斯) 正整数 n 不为三平方和当且仅当 n 可表成 $4^a(8b+7)$ 的形式,其中 a 和 b 为任意非负整数.

这个定理的一方面是容易的,如果 $n=4^a(8b+7)$ 是三平方和,$n=x^2+y^2+z^2$,易知 x,y,z 均为偶数(考查此式模 8 剩余).于是 $4^{a-1}(8b+7)$ 也为三平方和.这样下去可知 $8b+7$ 为三平方和.但是模 8 同余于 7 的数无法表成三平方和.这就表明 $4^a(8b+7)$ 不为三平方和.定理的另一半要困难得多.我们略去不讲,因为需要更多的数论知识.

根据三平方和的结果,拉格朗日(就是第 7 节所讲的热尔曼的老师)证明了:每个正整数都可表成四平方和,所以对每个 $k\geqslant 4$,也都可表成 k 个整数的平方和.

人们进一步要问:正整数表成 k 个整数的平方和有多少种方法?也就是说:方程

$$x_1^2+x_2^2+\cdots+x_k^2=n$$

有多少种整数解$(x_1,\cdots,x_k)$?我们把这个解数记成 $N_k(n)$.引理 10.1 给出了 $N_2(n)$ 的简单公式.高斯也给出了 $N_3(n)$ 的公式,但公式

中要用二次域(或者更正确地说是用二元二次型)的类数,所以 $N_3(n)$和 $N_2(n)$的公式形式和证明都有很大区别.而 $N_4(n)$的公式则与 $N_2(n)$相仿,即只与 n 的正因子有关.确切地说:$\frac{1}{8}N_4(n)$等于 n 的所有不被 4 除尽的正因子之和,即

$$N_4(n) = 8\sum_{\substack{d \mid n \\ 4 \nmid d}} d.$$

例如对 $n=12$, n 的所有不被 4 除尽的正因子为 1,2,3,6.所以方程 $x_1^2+x_2^2+x_3^2+x_4^2=12$ 的正整数解共有 $8(1+2+3+6)=96$ 个.事实上,这 96 个解为

$$(x_1, x_2, x_3, x_4)$$

$$= (0, \pm 2, \pm 2, \pm 2), (\pm 1, \pm 1, \pm 1, \pm 3)$$

以及它们的不同次序置换.1828 年,雅可比利用模函数得到 $N_6(n)$和 $N_8(n)$的公式.1864 年和 1866 年刘维尔分别得到 $N_{10}(n)$和 $N_{12}(n)$的公式.到 1916 年,对于偶数 k 人们算到 $N_{24}(n)$.而当 n 是奇数的情形,只有 $N_3(n)$、$N_5(n)$和$N_7(n)$是 19 世纪得到的,对于 $k=9$, $11, 13, \cdots, 23$, $N_k(n)$是在 1949 年才有公式.人们发现,当 k 是奇数时计算 $N_k(n)$比 k 为偶数

时要困难得多，而且公式的形式也差别很大．为了理解这件事，需要介绍计算 $N_k(n)$ 的方法，它是用一种解析工具，这种工具称为模形式．

让我们考虑复变函数

$$\theta(z)=\sum_{n=-\infty}^{\infty}e^{2\pi izn^2}=1+2\sum_{n=1}^{\infty}e^{2\pi in^2z}. \tag{10.5}$$

设 $z=\sigma+it$，其中 σ 和 t 为实数，分别称为复变量 z 的实部分和虚部分．由于 $|e^{2\pi in^2z}|=e^{-2\pi tn^2}$，可知当 $t>0$ 的时候，(10.5)式是(绝对)收敛的级数．所以函数 $\theta(z)$ 在区域

$$H=\{z=\sigma+it\mid t>0\}$$

中定义成解析函数．区域 H 是复平面在水平轴上面的那一半，称为上半平面．如果我们把 $k(\geqslant 1)$ 个 $\theta(z)$ 乘起来，便得到

$$\begin{aligned}\theta^k(z)&=\Big(\sum_{n_1=-\infty}^{\infty}e^{2\pi in_1^2z}\Big)\Big(\sum_{n_2=-\infty}^{\infty}e^{2\pi in_2^2z}\Big)\cdots\Big(\sum_{n_k=-\infty}^{\infty}e^{2\pi in_k^2z}\Big)\\&=\sum_{n_1,\cdots,n_k=-\infty}^{\infty}e^{2\pi i(n_1^2+n_2^2+\cdots+n_k^2)z}\\&=\sum_{n=0}^{\infty}N_k(n)e^{2\pi inz}\quad(z\in H)\end{aligned} \tag{10.6}$$

最后一个等式是因为：对每个 $n \geqslant 0$，当 $n_1, \cdots, n_k$ 取所有整数时，其中满足 $n_1^2 + n_2^2 + \cdots + n_k^2 = n$ 的恰好是 $N_k(n)$ 个．于是我们把所有要计算的值 $N_k(n)(n = 0, 1, 2, \cdots)$，归结于一个复变函数 $\theta^k(z)$ 的傅里叶展开系数．我们的问题是：如何通过研究函数 $\theta^k(z)$ 的解析特性来得到系数 $N_k(n)$ 的计算公式？

首先，由 $\theta(z)$ 的定义公式(10.5)可知 $\theta(z)$ 是周期为 1 的函数，所以对每个正整数 k，均有

$$\theta^k(z+1) = \theta^k(z) \quad (z \in H). \tag{10.7}$$

再下去人们发现了一个不平凡的函数关系：当 k 是 4 的倍数时($k = 4t, t = 1, 2, 3, \cdots$)，

$$\theta^{4t}\left(\frac{z}{4z+1}\right) = (4z+1)^{2t}\theta^{4t}(z) \quad (z \in H). \tag{10.8}$$

本世纪初，德国数学家赫克(Hecke)看出了这两个关系的深刻含义．我们要用到第 6 节最后介绍的群 G．G 中元素是方阵

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

其中 a, b, c, d 都是整数，并且 M 的行列式 $ad - bc$ 为 1．定义方阵 M 对复数 z 的作用为

$$M(z)=\frac{az+b}{cz+d}.$$

如果 $z=\sigma+it$ 在上半平面 H 中,即 $t>0$,则

$$\frac{az+b}{cz+d}=\frac{(a\sigma+b)+iat}{(c\sigma+d)+ict}$$

$$=\frac{(a\sigma+b+iat)(c\sigma+d-ict)}{(c\sigma+d)^2+t^2c^2}.$$

它的虚部分为

$$\frac{t}{(c\sigma+d)^2+t^2c^2}(ac\sigma+ad-ac\sigma-bc)$$

$$=\frac{t}{(c\sigma+d)^2+t^2c^2}>0,$$

所以 $M(z)$ 也属于上半平面 H. 于是 M 把上半平面 H 变成自身. 例如 $M_1=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ 和 $M_2=\begin{pmatrix}1&0\\4&1\end{pmatrix}$ 都属于 G,而由(10.7)和(10.8)式可知当 $M=\begin{pmatrix}ab\\cd\end{pmatrix}$ 取成 M_1 和 M_2 时,

$$\theta^{4t}(M(z))=(cz+d)^{2t}\theta^{4t}(z) \qquad (10.9)$$

是正确的. M_1 和 M_2 是群 G 中满足条件 $c\equiv 0\pmod 4$ 的方阵. 所有群 G 中满足 $c\equiv 0\pmod 4$ 的方阵 $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ 对于乘法也形成群,称为 G 的

一个子群.这个子群通常表成 $\Gamma_0(4)$,即

$$\Gamma_0(4)=\left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in G \mid c\equiv 0(\bmod 4)\right\}.$$

由于 M 取 M_1 和 M_2 时(10.9)式成立,可以推出(10.9)式对于子群 $\Gamma_0(4)$的任何方阵 M 都成立(用群论的语言说:这是由于 $\Gamma_0(4)$是由 M_1 和 M_2 生成的).换句话说,函数 $\theta^{4t}(z)$在 G 的一个子群 $\Gamma_0(4)$的作用下都有关系式(10.9),这样的函数称为模形式.也就是说,

定义 设 Γ 是 G 的一个子群,$f(z)$为上半平面 H 中的解析函数.如果对 Γ 中每个方阵 $M=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$,均有

$$f\left(\frac{az+b}{cz+d}\right)=(cz+d)^k f(z)\quad (z\in H),$$

我们称 $f(z)$为对于群 Γ 的权 k **模形式**.

于是,$\theta^{4t}(z)$是对于群 $\Gamma_0(4)$的权 $2t$ 的模形式,那么,$\theta^{4t}(z)$是模形式这件事对于我们计算 $\theta^{4t}(z)$的系数 $N_{4t}(n)$有什么好处呢? 20 世纪 30 年代以来,模形式发展了深刻的理论,形成了近代数论的一个重要分支.确切地说,这是数论、几何与函数论相交叉而形成的一个数学分支.我们只介绍其中一部分内容.

不难由定义看出,如果 f 和 g 都是对于群

Γ 的权 k 模形式，则对任意复数 α 和 β，函数 $\alpha f+\beta g$ 也是这样的模形式. 如果把所有对于群 Γ 的权 k 模形式组成的集合表示成 $M_k(\Gamma)$，则上面的论断是说：$M_k(\Gamma)$ 是复数域上的向量空间. 模形式理论的一个重要结果是说：这个向量空间是有限维的. 通俗一点说，就是 $M_k(\Gamma)$ 中存在着有限多个模形式 $f_1(z),\cdots,f_d(z)$（d 称为 $M_k(\Gamma)$ 的维数），使得 $M_k(\Gamma)$ 中每个模形式都惟一表达成

$$f(z)=\alpha_1 f_1(z)+\cdots+\alpha_d f_d(z),$$

其中 $\alpha_1,\cdots,\alpha_d$ 是复数. 我们称 $f_1,\cdots,f_d$ 为模形式空间 $M_k(\Gamma)$ 的一组基.

模形式理论最基本也是最重要的问题是：对于 G 的某个子群 Γ 和整数 $k\geqslant 2$，

（1）计算对于群 Γ 的权 k 模形式空间 $M_k(\Gamma)$ 的维数 d；

（2）寻求 $M_k(\Gamma)$ 的一组基 $f_1,f_2,\cdots,f_d$. 到目前为止，只是对一部分群 Γ 和权 k 解决了上述问题. 而对许多情形，不但没有找到一组基，甚至连维数 d 也没有计算出来.

现在我们举例说明模形式理论如何用来计算 $N_k(n)$ 的值. 取 $k=4$，我们知道

$$\theta^4(z)=\sum_{n=0}^{\infty} N_4(n)e^{2\pi i n z}\quad (z\in H)$$

是对于群 $\Gamma_0(4)$ 的权 2 模形式，即属于 $M_2(\Gamma_0(4))$. 人们算出这个空间的维数是 2，并且也找到了 2 维空间 $M_2(\Gamma_0(4))$ 的一组基 $f_1(z)$ 和 $f_2(z)$，其中

$$f_1(z) = \sum_{2 \nmid n} \sigma(n) e^{2\pi i n z} \quad (z \in H),$$

其中求和表示 n 过所有正奇整数，而 $\sigma(n)$ 表示 n 的所有正整数因子之和（例如 $\sigma(15) = 1+3+5+15=24$）.

$$f_2(z) = \frac{1}{24} + \sum_{n=1}^{\infty} \sigma'(n) e^{2\pi i n z} \quad (z \in H),$$

其中 $\sigma'(n)$ 表示 n 的所有正奇整数因子之和（例如 $\sigma'(20) = 1+5=6$）. 于是 $\theta^4(z)$ 可以惟一表达成

$$\theta^4(z) = \alpha f_1(z) + \beta f_2(z).$$

比较这三个函数的系数，可知对所有正整数 $n \geqslant 1$，

$$N_4(n) = \begin{cases} \alpha\sigma(n) + \beta\sigma'(n), & \text{若 } n \text{ 为奇数}, \\ \beta\sigma'(n), & \text{若 } n \text{ 为偶数}. \end{cases}$$

但是 α 和 β 可以由两个 n 值决定出来：取 $n=1$，则 $\sigma'(n) = \sigma(n) = 1$，而 $N_4(1) = 8$（因为 $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$ 有 8 个整数解）. 所以

$$8=\alpha+\beta.$$

再取 $n=2$，则 $\sigma'(2)=1$，而 $N_4(2)=24$（容易算出 $x_1^2+x_2^2+x_3^2+x_4^2=2$ 共有 24 个整数解）. 所以

$$24=\beta.$$

由此得到 $\alpha=-16$，$\beta=24$ 这就给出 $N_4(n)$的公式

$$N_4(n)=\begin{cases}24\sigma'(n)-16\sigma(n), & \text{若 } n \text{ 为正奇数},\\ 24\sigma'(n), & \text{若 } n \text{ 为正偶数}.\end{cases} \tag{10.10}$$

现在令 $\tau(n)$表示 n 的所有不被 4 除尽的正整数因子之和，我们在前面说 $N_4(n)=8\tau(n)$. 这与公式(10.10)是一致的，因为若 n 为正奇数，则$\sigma(n)=\sigma'(n)=\tau(n)$，于是 $N_4(n)=24\tau(n)-16\tau(n)=8\tau(n)$，而当 n 为正偶数时，n 的每个奇因子 r 对应着两个不被 4 除尽的因子 r 和 $2r$(其和为 $3r$)，所以 $\tau(n)=3\sigma'(n)$. 于是也有 $N_4(n)=24\sigma'(n)=8\tau(n)$.

对每个 $k=4t(t=1,2,3,\cdots)$，利用模形式理论原则上都可求出 $M_{2t}(\Gamma_0(4))$的一组基，所以用上述方法都能给出 $N_{4t}(n)$的公式，不过当 t 很大时，公式愈来愈复杂，不像 $N_4(n)$那

样简洁.对于每个 $k=2t(t=1,3,5,\cdots)$,也可算出 $M_t(\Gamma_0(4))$的维数并且寻求出一组基,但是比 $k=4t$ 情形要困难一些.当 k 为奇数时,则需要研究所谓"半整权"(即权 $k/2$ 是半整数)的模形式.半整权模形式理论从 20 世纪 70 年代之后才有重大进展,日本数学家志村五郎作了开创性的工作(这位数学家我们在第 12 节还会提到).这就是为什么对于奇整数 k,寻求计算 $N_k(n)$的公式是非常困难的.

我们不打算介绍模形式理论的进一步结果.但是要叙述一下赫克在 20 世纪中期得到的一个非常重要的结论,这个结论与费马猜想的后期发展有关联.对每个正整数 $N\equiv 0\pmod 4$,我们以 $\Gamma_0(N)$表示群 G 中满足 $c\equiv 0\pmod N$的方阵$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$全体.这也是 G 的一个子群.

定理 10.3(赫克) 设 $k\geqslant 2$ 为正整数,$M_k(\Gamma_0(N))$是对于群 $\Gamma_0(N)$的权 k 模形式空间,d 是此空间的维数.则此空间一定存在一组基 $f_1,\cdots,f_d$,它们都具有非常好的解析特性,现在被称为赫克模形式.

让我们来解释一下什么叫赫克模形式,它有什么样的解析性质.将模形式 $f(z)$写成傅里叶展开形式

$$f(z) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z} \quad (z \in H).$$

(10.11)

我们可以相应地写出一个级数

$$L_f(s) = \sum_{n=1}^{\infty} a_n n^{-s} \qquad (10.12)$$

称为模形式 $f(z)$的 L 函数. $f(z)$称为赫克模形式,是指它的 L 函数有以下一些好的解析性质.

(1) $L_f(s)$可以解析延拓成整个 s 复平面上的解析函数(而级数(11.10)本身只对 s 的实部分充分大时才收敛).并且有函数方程

$$L_f(s) = F(s) L_f(k-s)$$

其中 k 是模形式的权,而 $F(s)$是某个具体函数(我们没有把它明确写下来).

(2) $L_f(s)$具有以下形式的欧拉无穷乘积展开

$$L_f(s) = \prod_p \frac{1}{(1 - a_p p^{-s} + p^{k-2s})},$$

(10.13)

其中 k 为模形式 $f(z)$的权,当 p 为素数时,(10.13)式中的 a_p 和(10.11)、(10.12)式中的

系数 a_p 是一样的.

定理 10.3 是说:许多模形式空间都可以找到一组赫克模形式为基.这是一个很有用的定理.因为它们的 L 函数可以解析开拓!并且有函数方程和欧拉无穷乘积展开.从前几节的例子我们知道,这些解析性质对于研究数论都是很重要的.让我们举一个简单的例子.仔细比较一下 $L_f(s)$的两个表示式(10.12)和(10.13).在(10.12)式中出现所有的系数 a_n($n=0,1,2,\cdots$),它们也是模形式 $f(z)$的系数,而这些系数常常有某种数论意义.但是在(10.13)式中却只出现 a_p(下标 p 只过素数).这就意味着:所有的系数 a_n($n=1,2,3,\cdots$)应当由其一部分 a_p(p 过素数)所完全决定.事实上,我们把(10.13)中的级数展开然后比较(10.12)式可以得出如下的关系.

(1) $a_1=1$,对于 $n\geqslant 2$, 如果 $n=p_1^{c_1}\cdots p_s^{c_s}$,其中 $p_1,\cdots,p_s$ 是不同的素数,则

$$a_n = a_{p_1^{c_1}}\cdots a_{p_s^{c_s}}.$$

(2) 设 p 为素数而 $c\geqslant 2$,则

$$a_{p^c} = a_{p^{c-1}}a_p - pa_{p^{c-2}}.$$

这就表明所有 a_n 可由 a_p(p 为素数)所决定.

赫克定理是模形式理论中一个漂亮的结

果.但是在定理发现之后的半个世纪里,人们绝没有想到它会引发出费马猜想的最终解决.这中间有一个过程,即在20世纪50到70年代,由赫克定理猜想出模形式和椭圆曲线的深层联系.所以我们要步入到数论的另一个领域:椭圆曲线的算术.

11 椭圆曲线(1):有理点群

据考证,在公元 984 年,阿拉伯人就研究过如下的数学问题:

哪些正整数是有理直角三角形的面积?

所谓有理直角三角形,是指直角三角形的三边长度都是有理数.历史上把这样的正整数 n 称为"同余数"(congruent number).也就是说,n 是同余数,是指存在正有理数 a,b,c,使得

$$a^2+b^2=c^2,\quad n=\frac{1}{2}ab.$$

例如,$(a,b,c)=(3,4,5)$是直角三角形的三个边,所以$\frac{3\cdot 4}{2}=6$ 是同余数.又如取$(a,b,c)=\left(\frac{3}{2},\frac{20}{3},\frac{41}{6}\right)$,可知$\frac{1}{2}\cdot\frac{3}{2}\cdot\frac{20}{3}=5$ 是同余数.另一方面,1,2,3 都不是同余数,据说这也是费马的结果.我们来证明:

1 不是同余数.

证 如果 1 是同余数,则存在正有理数 a,

b, c,使得 $a^2+b^2=c^2$, $ab=2$.于是

$$(a+b)^2=c^2+4, (a-b)^2=c^2-4.$$

$$(a+b)^4-(a-b)^4=(c^2+4)^2-(c^2-4)^2$$
$$=16=4^2.$$

这表明方程 $x^4-y^4=z^2$ 存在有理数解$(x, y, z)=(a+b, |a-b|, 4)$.由 $a^2+b^2=c^2$ 易知 $a\neq b$(因为若 $a=b$,则 $2a^2=c^2$,推出$\sqrt{2}=\frac{c}{a}$为有理数),所以上面给出的是正有理数解.进而,若(x, y, z)满足上面方程,则对每个整数 m, (mx, my, m^2z)也满足此方程.所以由方程的一组正有理数解一定如此法得到一组正整数解.但是用费马的无穷下降法可以证明方程 $x^4-y^4=z^2$没有正整数解(这是第 5 节的习题).这一矛盾表明 1 不是同余数.

习题:设 n 和 m 均为正整数.则 n 和 nm^2 或者均为同余数,或者均不是同余数(提示:将直角三角形按 $m:1$ 的比例放大或缩小).

根据这个习题,我们以后只需考虑 n 是无平方因子的正整数.

图 5 是一个面积为 157 的有理直角三角形,用来证明 157 是同余数.三个边长的分母都很大,但这是所有面积 157 的直角三角形当中分母最小的一个!

同余数问题看起来很容易,但实际上非常

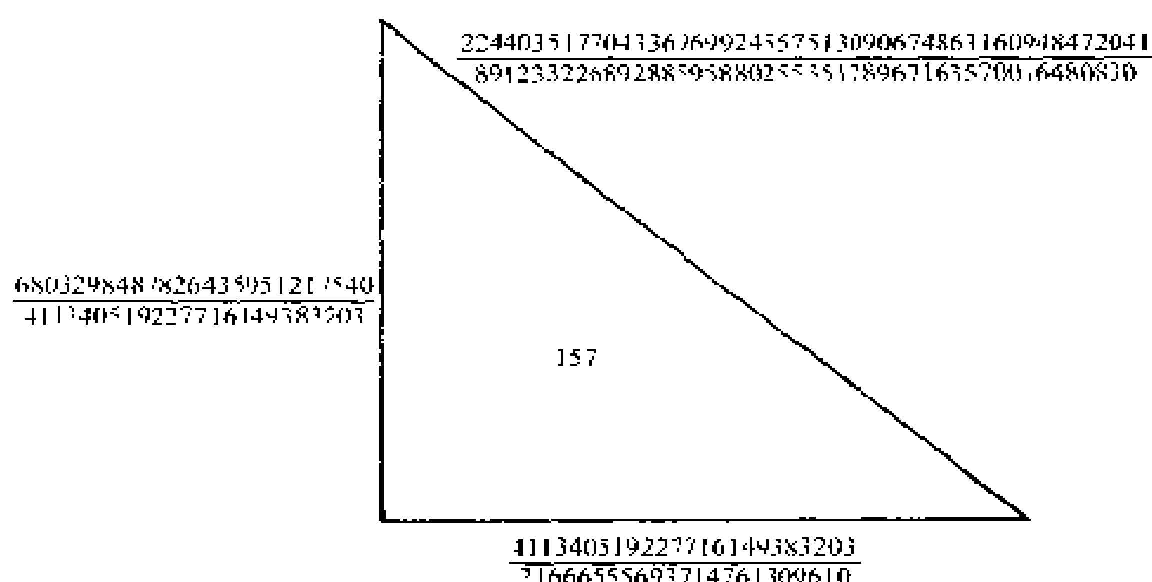


图 5　157 是同余数

难. 事实上这个问题至今未完全解决. 到目前为止, 已经证明了所有素数 $p\equiv 5,7 \pmod 8$ 都是同余数, 所有素数 $p\equiv 3 \pmod 8$ 都不是同余数. 1960 年, 有三位数学家猜想: 所有正整数 $n\equiv 5,6,7 \pmod 8$ 都是同余数, 这个猜想至今没有解决. 近年来关于同余数问题的新进展都是和椭圆曲线的最新进展有密切关系. 它的出发点是人们早就认识到的如下事实:

n 为同余数当且仅当椭圆曲线 $y^2=x^3-n^2x$ 有非零有理数解 (x,y) (即 x 和 y 均是非零有理数).

在证明这个事实之前, 我们要介绍椭圆曲线的一个基本但是很重要的结果, 就是: 椭圆曲线上的全部有理点组成的集合可以定义一个加法运算, 使得这个集合形成一个交换群. 这个加

法运算非常特别，并且是采用的几何办法.

椭圆曲线 E 的典型方程为

$$E: y^2 = x^3 + ax + b, \qquad (11.1)$$

其中 a 和 b 为整数，并且 $4a^3 + 27b^2 \neq 0$（后一条件意味着多项式 $x^3 + ax + b$ 在复数域中没有重根）.

方程(11.1)的全部实数解画在坐标平面上是一条（或两条）曲线. 由于 $x^3 + ax + b$ 为实系数方程，熟知它有一个或三个实根. 对 $x^3 + ax + b$ 的每个实根 α，曲线(11.1)有实数点 $(x, y) = (\alpha, 0)$ 在 x 轴上. 所以椭圆曲线 E 与 x 轴交于 1 点或 3 点. 若 $P = (x, y)$ 是 E 上一点，则 $\bar{P} = (x, -y)$ 也是 E 上一点. 所以椭圆曲线 E 的图形关于 x 轴是对称的. 当 $x^3 + ax + b$ 有三个实根 $\alpha < \beta < \gamma$ 时，椭圆曲线 E 的图形大致如图 6 所示

我们先从直观几何上说明在椭圆曲线 E 上如何定义加法. 设 $P = (x_1, y_1)$ 和 $Q = (x_2, y_2)$ 是曲线 E 上的两个不同的实数点. 过点 P 和 Q 的直线 l 与曲线 E 交于第 3 点 $R = (x_3, y_3)$，我们把与 R 对称的点 $\bar{R} = (x_3, -y_3)$ 定义为点 P 和点 Q 的“和”，表示成

$$P \oplus Q = \bar{R}.$$

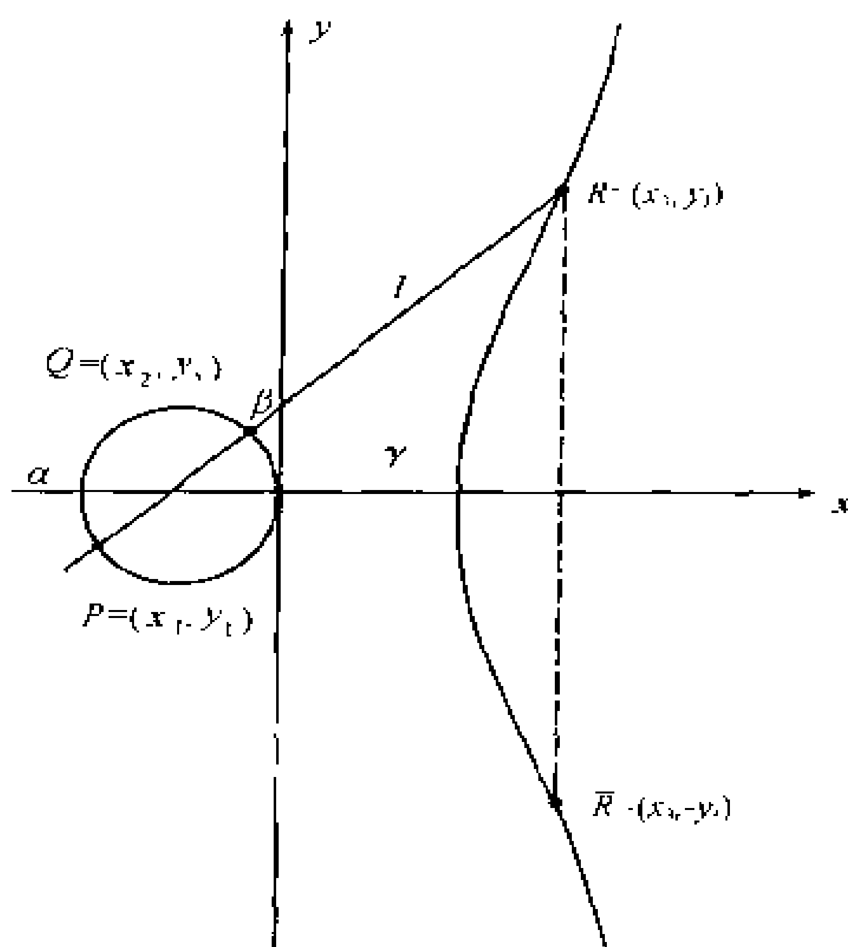


图 6　椭圆曲线的加法

注意这种加法满足交换律，因为直线 l 由点 P 和 Q 所决定，与 P 和 Q 的次序无关，于是

$$P \oplus Q = Q \oplus P.$$

这样的几何直观仔细推敲起来，还有许多问题：

(1) 直线 l 上除了点 P 和 Q 之外，如果与曲线 E 交于两个点怎么办？我们要证明这是不可能的．另一个情形是：如果直线 l 只与 E 有

两个公共点(没有第三个交点 R)怎么办?这是完全可能的.比如说取曲线上两个对称点 $P=(\alpha,\beta)$和 $\bar{P}=(\alpha,-\beta)(\beta\neq 0)$.则连接 P 和 P 的直线 l 是与 y 轴平行的直线 $x=\alpha$.由于 $x=\alpha$时,方程 $y^2=\alpha^3+a\alpha+b$ 只能有两个解 $y=\pm\beta=\pm\sqrt{\alpha^3+a\alpha+b}$,所以直线 $x=\alpha$ 不能与曲线 E 再交于第三点.为了克服这个困难,我们人为地假设椭圆曲线 E 上还有一个“无穷远点”∞,这个点也在所有与 y 轴平行的直线上.这样一来,直线 $x=\alpha$ 和椭圆曲线 E 交于第三点 ∞.我们还假定 $\infty=\bar{\infty}$,于是 $P+\bar{P}=\bar{\infty}=\infty$.现在我们把 ∞ 和 P 相加,过 ∞ 和 P 的直线应当是过 P 并且与 y 轴平行的直线(因为 ∞ 只在与 y 轴平行的直线上),这条直线与 E 的另一个交点为 $\bar{P}$.按照定义,$\infty\oplus P$ 应当是 $\bar{P}$ 的对称点 P.于是我们就得到:对椭圆曲线 E 上的每个点 P,均有 $\infty\oplus P=P$.所以无穷远点对于我们定义的加法运算实际上起着零元素的作用.因此我们今后把 ∞ 改记成 0.于是 $P+\bar{P}=0$,从而 $\bar{P}=-P$.

(2) 我们对 E 上两个不同的点 P 和 Q,可以连一条直线 l.如果 $P=Q$ 怎么办?即如何定义 P 和 P 相加(表示成[2]P)?这种情形则要用“连续性法则”,即我们在 P 的附近取 E 上一个点 $Q(Q\neq P)$.连结 P 和 Q 的直线 l 与曲

线 E 交于第3点 R，于是 $P\oplus Q=\bar{R}$. 现在把 Q 趋近于 P，熟知直线 l 变成曲线 E 上过点 P 的切线. 于是 R 趋近于这条切线与 E 的交点 S. 而 $\bar{R}$ 趋近于 $\bar{S}$. 所以我们定义 $[2]P=P\oplus P=\bar{S}$.

(3) 根据上面的定义，我们看到这种运算有零元素 0，而且每个点 P 有逆元素 $-P=\bar{P}$. 可是要想使曲线 E 上的全部实数点(连同 0)对此运算形成一个群，还需要一个最起码的要求，就是结合律：$(P\oplus Q)\oplus R=P\oplus(Q\oplus R)$. 这件事从几何图形上是很难看出来的，但它是对的. 我们需要把由几何直观想象出来的加法运算严格化和代数化. 即若 $P=(x_1,y_1)$，$Q=(x_2,y_2)$，$P\oplus Q=\bar{R}=(x_3,-y_3)$，我们希望给出用坐标 (x_1,y_1) 和 (x_2,y_2) 表达 x_3 和 y_3 的代数公式.

如前所设，椭圆曲线 E 的方程为

$$E: y^2=x^3+ax+b$$

$$(a,b\in\ ,\ 4a^3+27b\neq 0).\quad (11.1)$$

(1) $P=(x_1,y_1)$ 和 $Q=(x_2,y_2)$ 是椭圆曲线 E 上两个实数点. 并且 $x_1\neq x_2$. 这时连结 P 和 Q 的直线 l 不与 y 轴平行，所以有方程

$$l: y=kx+t,$$

其中

$$k=\frac{y_2-y_1}{x_2-x_1},t=y_1-kx_1(=y_2-kx_2).$$

$$(11.2)$$

直线 l 与曲线 E 的交点是方程组

$$\begin{cases}y=kx+t,\\ y^2=x^3+ax+b\end{cases}\quad(11.3)$$

的公共解.将第一方程代人第二方程,得到

$$x^3+ax+b-(kx+t)^2=0.\quad(11.4)$$

这是关于 x 的实系数三次方程.由于 $P=(x_1,y_1)$和 $Q=(x_2,y_2)$是方程组(11.3)的公共解,可知 x_1,x_2 是方程(11.4)的两个实根.所以第三个根也必为实根.事实上,由于(11.4)式左边 x^2 的系数为 $-k^2$,由韦达定理可知方程(11.4)的第三个根 x_3 满足 $x_1+x_2+x_3=k^2$,即

$$x_3=k^2-x_1-x_2=\left(\frac{y_2-y_1}{x_2-x_1}\right)^2-x_1-x_2.$$

$$(11.5)$$

于是直线 l 一定与 E 交于另一点 $R=(x_3,y_3)$,其中

$$y_3 = kx_3 + t = kx_3 + y_1 - kx_1$$

$$= \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x_3 - x_1) + y_1. \quad (11.6)$$

这就表明：当 $x_1 \neq x_2$ 时，$P \oplus Q = (x_3, -y_3)$，其中 x_3 和 y_3 由公式(11.5)和(11.6)给出.

(2) 设 $x_1 = x_2$，则在 $P \neq Q$ 时必然 $Q = P$，定义 $P + P = 0$.

(3) 设 $P = (x_1, y_1)$. 我们推导 $[2]P = (x_3, -y_3)$ 的公式. 为此，我们取 ε 为充分小的正实数，则 $Q = (x_2, y_2)$ 是与 P 很近的 E 上一点，其中

$$x_2 = x_1 + \varepsilon, y_2^2 = x_2^3 + ax_2 + b$$

$$= (x_1 + \varepsilon)^3 + a(x_1 + \varepsilon) + b.$$

令 $\varepsilon \to 0$，则连结 P 和 Q 的直线变成曲线 E 过点 P 的切线. 这个切线的斜线为 $k = \frac{y_2 - y_1}{x_2 - x_1}$ 的极限：

$$\lim_{\varepsilon \to 0} \frac{y_2 - y_1}{x_2 - x_1} = \lim_{\varepsilon \to 0} \frac{y_2^2 - y_1^2}{\varepsilon(y_2 + y_1)}$$

$$= \lim_{\varepsilon \to 0} \frac{(x_1 + \varepsilon)^3 + a(x_1 + \varepsilon) + b - x_1^3 - ax_1 - b}{2\varepsilon y_1}$$

不难算出这个极限值为$\frac{3x_1^2+a}{2y_1}$,所以对(11.5)和(11.6)式取极限就得到

$$x_3=\left(\frac{3x_1^2+a}{2y_1}\right)^2-2x_1, \qquad (11.7)$$

$$y_3=\frac{3x_1^2+a}{2y_1}(x_3-x_1)+y_1. \ (11.8)$$

这就表明对于 E 上点 $P=(x_1,y_1)$,当 $y_1\neq 0$ 时,$[2]P=(x_3,-y_3)$,其中 x_3 和 y_3 由公式(11.7)和(11.8)算出.而当 $y_1=0$ 时,$P=\overline{P}=-P$,因此$[2]P=0$.

以上我们对于椭圆曲线 E 上的实数点(以及 0)严格得到加法运算的坐标表达式.可以用这些表达式来验证加法运算满足结合律(和交换律),验证需要很麻烦的运算,我们这里从略.综合上述,我们把椭圆曲线 E 上所有的实数点连同 0 一起组成的集合表示成$E(*)$,则 $E(*)$ 对于上述定义的加法运算是一个交换群,零元素为 0,而 $-(x,y)=(x,-y)$.

在数论中我们关心的是椭圆曲线 E 上的有理数点.如果 $P=(x_1,y_2)$和 $Q=(x_2,y_2)$为 E 上的有理数点.由于 a 和 b 为整数,可知由(11.5)~(11.8)算出的点(x_3,y_3)也是有理数点.这表明在 E 的所有有理数点(以及 0)组成

的集合 $E(\ )$ 上也可进行上述加法运算，并且 $E(\ )$ 对这个运算为交换群.

现在我们举两个例子：

例 1 考虑椭圆曲线 $E: y^2 = x^3 - 432$. 可直接验证 $P = (12,36)$ 是 E 上的有理数点. 用公式(11.7)和(11.8)算出 $[2]P = (12, -36) = P$. 于是 $[3]P = \overline{P} + P = 0$.

例 2 椭圆曲线 $E: y^2 = x^3 - 36x$ 上有点 $P = (-3,9)$ 和 $Q = (-2,8)$. 用公式可算出新的有理数点

$$P \oplus Q = (6,0), \quad [2]P = \left(\frac{25}{4}, -\frac{35}{8}\right).$$

于是 $[2]P + [2]Q = [2](P \oplus Q) = 0$, 从而 $[2]Q = -[2]P = \left(\frac{25}{4}, \frac{35}{8}\right)$. 可以证明 $\pm P, \pm [2]P, \cdots, \pm [n]P, \cdots$ 是彼此不同的点，所以 E 上有无穷多个有理数点，即方程 $y^2 = x^3 - 36x$ 有无穷多个有理数解.

从以上两个例子知道，椭圆曲线 E 上的有理数点 P 可能有以上两种情形：

(1) 存在正整数 n, 使得 $[n]P = 0$. 满足此式的最小正整数 n 称为点 P 的**阶**，而 P 称为**有限阶点**. 由点 P 只能得到有限多个有理数点：$P, [2]P, \cdots, [n-1]P[n]P = 0$. (而 $[n+1]P = P$, $-P = [n-1]P$.)

(2) 不存在正整数 n,使得$[n]P=0$.这时 P 称为**无限阶**元素,由 P 可得到 E 上无限多个有理数点$[n]P(n=0,1,2,\cdots)$.

我们希望通过加法运算能够有效地得到椭圆曲线 E 的全部有理数点,1925 年,英国数学家莫德尔(Mordell)得到了一个重要的理论结果:由椭圆曲线 E 上**有限个**有理数点通过加法运算就可以得到全部有理数点.确切地说,我们有

定理 11.1(莫德尔) 对每个椭圆曲线 E,$E(\ )$是有限生成交换群.也就是说:

(1) 椭圆曲线 E 上的有限阶有理数点只有有限多个.所有的有限阶有理数点组成的集合 $E(\ )_t$ 是 $E(\ )$的子群,并且 $E(\ )_t$ 是有限群.

(2) 存在整数 $r=r(E)\geqslant 0$ 和曲线 E 上 r 个无限阶有理数点 $P_1,P_2,\cdots,P_r$,使得 E 上每个有理数点都惟一地表示成

$$P+[n_1]P_1+\cdots+[n_r]P_r, \quad (11.9)$$

其中 $P\in E(\ )_t$,而 $n_1,\cdots,n_r$ 为整数.

定理中的 $P_1,\cdots,P_r$ 称为有理数点群 $E(\ )$的一组**基**,而 r 称为椭圆曲线 E 的**秩**.如果 $r=0$,则 $E(\ )=E(\ )_t$ 是有限群,即 E 上只有有限个有理数点.如果 $r\geqslant 1$,则 $E(\ )$中

存在无限阶点，所以 E 上有无限多个有理数点. r 愈多则 E 上的有理数点愈多.

这样一来，椭圆曲线上的算术理论主要研究有理点群 $E(\mathbb{Q})$ 的结构. 基本问题有：

(1) 决定 E 上所有的有限阶有理数点（这只有有限多个）. 换句话说，要决定有限交换群 $E(\mathbb{Q})_t$.

(2) 决定 E 的秩 $r = r(E)$.

(3) 决定 $E(\mathbb{Q})$ 的一组基 $P_1, \cdots, P_r$.

问题(1)是容易的，目前已有好的方法决定 E 上全部有限阶有理数点. 比如说，可以证明椭圆曲线 $E: y^2 = x^3 - n^2 x$ 上只有 4 个有限阶有理点：$(x, y) = (0, 0)$, $(\pm n, 0)$ 和 O，其中 O 为 1 阶点，而 $(0,0)$ 和 $(\pm n, 0)$ 均是 2 阶点. 所以 E 上的非零有理数点一定是无限阶的. 另一方面，决定 $r = r(E)$（问题(2)）是相当困难的，甚至于决定 r 是否为 0（即 E 是否只有有限多个有理数点）也是相当困难的. 现在人们猜想：对每个正整数 N 都存在椭圆曲线 E 使得 $r(E) > N$. 这个猜想至今未能解决. 目前用深刻数论技巧和使用计算机，所发现的椭圆曲线的最大秩为 14，而问题(3)则更加困难.

作为本节的一个应用，我们证明关于同余数和椭圆曲线关系的前述结果.

定理 11.2 对于正整数 n，下列三个条件

是彼此等价的

(1) n 为同余数.

(2) 椭圆曲线 $E_n: y^2 = x^3 - n^2 x$ 的秩 $r = r(E_n)$ 为正整数.

(3) 方程 $y^2 = x^3 - n^2 x$ 有非零有理数解.

证 由于 $y^2 = x^3 - n^2 x$ 的非零有理数解都是无限阶点,可知(2)和(3)是等价的.我们只需证(1)和(3)等价.设 n 是同余数,即存在正有理数 a, b, c,使得 $a^2 + b^2 = c^2$, $ab = 2n$.如果 $a = b$,则 $2a^2 = c^2$,推出 $\sqrt{2} = \frac{c}{a}$ 为有理数.因此 $a \neq b$.于是

$$A = \frac{c^2}{4}, \quad B = \frac{c(a^2 - b^2)}{8}$$

都是非零有理数.由于

$$\begin{aligned} A^3 - n^2 A &= \frac{c^2}{4}\left(\frac{c^4}{16} - n^2\right) \\ &= \frac{c^2}{4} \cdot \frac{(a^2 + b^2)^2 - 4a^2b^2}{16} \\ &= \left(\frac{c(a^2 - b^2)}{8}\right)^2 = B^2. \end{aligned}$$

可知 $(x, y) = (A, B)$ 是方程 $y^2 = x^3 - n^2 x$ 的非零有理数解.

现在设 $y^2 = x^3 - n^2 x$ 有非零有理数解

$(x,y)=(A,B)$. 我们来证 n 为同余数，由椭圆曲线 E_n 上的有理数点 $P=(A,B)$ 用公式 (11.7) 和 (11.8) 算出新的有理数点 $[2]P=(M,-N)$. 换句话说，曲线 E_n 在点 P 的切线与 E_n 交于点 (M,N). 设此切线为 $y=kx+t$，我们知道三个点 P,P 和 (M,N) 是切线与 E_n 的三个交点，即是方程组

$$\begin{cases} y = kx + t, \\ y^2 = x^3 - n^2 x \end{cases}$$

的全部解. 所以 $x^3-n^2x-(kx+t)^2$ 的三个根为 A,A 和 M，于是

$$x^3 - n^2 x - (kx + t)^2 = (x - A)^2 (x - M).$$

对于 x^3-n^2x 的三个根 $\alpha=0,\pm n$，将 $x=\alpha$ 代入上式得到

$$(k\alpha + t)^2 = (\alpha - A)^2 (M - \alpha) \quad (\alpha = 0, \pm n).$$

这就表明 $M-\alpha(\alpha=0,\pm n)$ 都是有理数的平方(由 $(x,y)=(A,B)$ 是 $y^2=x^3-n^2x$ 的解并且 $B\neq 0$，可知 $A-\alpha\neq 0$). 令

$$M = u^2, M + n = v^2, M - n = w^2, \tag{11.10}$$

其中 u, v, w 为有理数. 如果 $M-n=0$, 则 $M=n, n=u^2, 2n=v^2$, 推出 $\sqrt{2}$ 为有理数的平方. 这表明 $M-n>0$, 所以 M 和 $M+n$ 也大于零. 因此可认为 u, v, w 都是正有理数, 并且 $0<w<u<v$. 记

$$H=2u, F=v+w, G=v-w,$$

则 H, F, G 是正有理数. 由(11.10)可知

$$\begin{aligned} F^2+G^2 &= (v+w)^2+(v-w)^2 \\ &= 2(v^2+w^2)=4M=4u^2=H^2, \end{aligned}$$

$$\frac{1}{2}FG=\frac{1}{2}(v^2-w^2)=n.$$

这就表明 n 是同余数.

在整个 20 世纪里, 对于椭圆曲线有理点群的研究是数论的一个中心课题. 引进了几何、代数和解析手段, 并且彼此有紧密联系. 我们在下节介绍研究椭圆曲线的解析方法.

12 椭圆曲线(2):L 函数

我们仍设椭圆曲线 E 有方程

$$E:\ y^2 = x^3 + ax + b$$

$$(a, b \in \quad, 4a^3 + 27b^2 \neq 0).$$

上一节已经把曲线 E 上有理数点连同 O 做成一个交换群 $E(\quad)$,其全部有限阶点形成一个有限子群 $E(\quad)_t$,并且有一组基 P_1、…,P_r,使得每个有理数点都可由有限阶点与基中点通过运算得到.椭圆曲线 E 的秩 r,群 $E(\quad)_t$ 中元素个数,基 P_1,…,P_r 的性质等都是 E 的重要算术性质.所谓用解析方法研究椭圆曲线的算术,就是探讨是否对每个椭圆曲线 E 都能够找到一个适宜的复变函数,使得此函数的解析性质反映 E 的上述算术性质.

历史上,这种考虑首先不是在有理数域上进行,而是在有限域 $\ _p$ 上进行的.由于 a 和 b 是整数,所以对每个素数 p,都可以把 $y^2 = x^3 + ax + b$ 看成是有限域 $\ _p$ 上的方程,即同余

方程

$$y^2 \equiv x^3 + ax + b \pmod{p}.$$

如果 $4a^3+27b^2$ 在 $\mathbb{F}_p$ 中仍不为零，即 $4a^3+27b^2 \not\equiv 0 \pmod{p}$，我们也把 E 看成是 $\mathbb{F}_p$ 上一条椭圆曲线．由于已假定 $4a^3+27b^2$ 是非零整数，所以它的素因子只有有限个．于是除了这有限个素因子之外的每个素除 p，E 都是 $\mathbb{F}_p$ 上的椭圆曲线．我们可以像上节一样用公式(10.5)—(10.8)定义加法运算，使得 E 在 $\mathbb{F}_p$ 中的全部点(即 $y^2=x^3+ax+b$ 在 $\mathbb{F}_p$ 中的全部解)连同 O 成为交换群，这个群记成 $E(\mathbb{F}_p)$．与群 $E(\mathbb{Q})$ 不同，$E(\mathbb{F}_p)$ 一定是有限群，因为 $\mathbb{F}_p$ 上所有点 (x,y) 只有 p^2 个(每个 x,y 在 $\mathbb{F}_p$ 中都可选其中 p 个元素)，所以满足 $y^2=x^3+ax+b \pmod{p}$ 的也只有有限多个．我们以 N_p 表示有限群 $E(\mathbb{F}_p)$ 的元素个数．

20 世纪初期，人们就对有限域 $\mathbb{F}_p$ 上的椭圆曲线 E 构作了一个解析函数．1935 年，英国数学家达文波特(Davenport)和德国数学家哈塞(Hasse)证明了这个函数有类似于黎曼猜想的性质，即函数的根的实部分均为 $\frac{1}{2}$．由此给出有限群 $E(\mathbb{F}_p)$ 元素个数 N_p 的一个重要估计：

$$|N_p-(p+1)|\leqslant 2\sqrt{p}.\qquad(12.1)$$

所以当 $p\to\infty$ 时,$\frac{p}{N_p}$的极限为 1.

例 考虑椭圆曲线 $E: y^2=x^3-x$.由于 $a=-1, b=0, 4a^3+27b^2=-4$.于是对每个奇素数 p,E 是 $\mathbb{F}_p$ 上一条椭圆曲线,以 $p=7$ 为例,同余方程

$$y^3\equiv x^3-x\pmod 7$$

在 $\mathbb{F}_7$ 中共有 7 个解

$$(x,y)=(0,0),(\pm 1,0),(4,\pm 2)\text{和}(5,\pm 1).$$

连同无穷远点 O,可知 $E(\mathbb{F}_7)$是有 8 个元素的交换群.对于 $P=(4,2)$, $Q=(0,0)$,用加法公式在 $\mathbb{F}_p$ 中可算出

$$P\oplus Q=(5,-1),[2]P=(1,0),$$

可直接验证群 $E(\mathbb{F}_7)$的 8 个元素为

$$[n]P+[m]Q\quad(n=0,1,2,3, m=0,1).$$

于是 $N_7=8$,而$|N_7-(7+1)|=0\leqslant 2\sqrt{7}$.

有限域上的椭圆曲线和有理数域上的椭圆曲线应当有什么联系吗?人们有一种朴素的看法,如何椭圆曲线 E 在每个有限域 $\mathbb{F}_p$ 上都有较多的点(即 N_p 较大),那么 E 在有理数域上也

应当有较多的点(即群$E(\mathbb{Q})$较大).问题在于:这种看法是否有道理?而且既使有道理,如何把它表示成定量的形式和确切的关系?

由于当$p\to\infty$时$\frac{p}{N_p}$的极限是1,人们便考虑无穷乘积

$$\prod_p \frac{p}{N_p},$$

其中p过所有素数.这是无穷多个正实数相乘.所以其值$\geqslant 0$.如果N_p较大,则$\frac{p}{N_p}$较小,使这个无穷乘积达到最小值0,则人们期望E也有较多的有理数点,即期望有无穷多有理点.如果N_p较小,使无穷乘积的值>0,则期望E也有较少的有理数点,即期望只有有限多个有理数点.也就是说,上面朴素的想法引导出如下的定量性的猜测

$$\prod_p \frac{p}{N_p}=0 \Leftrightarrow E\text{ 有无穷多有理数点} \Leftrightarrow \text{群 } E(\mathbb{Q})\text{ 的秩 } r\geqslant 1,$$

$$\prod_p \frac{p}{N_p}>0 \Leftrightarrow E\text{ 只有有限多有理数点} \Leftrightarrow r=0.$$

1958～1960年,英国两位数学家Birch和

Swinnerton-Dyer对大量椭圆曲线进行具体计算，上面的猜测对这些计算例子都正确．在这种实践的基础上，他们构作出椭圆曲线 E 的一个复变函数

$$L(E,s) = \prod_p \frac{1}{1 - a_p p^{-s} + p^{1-2s}}, \tag{12.2}$$

其中 p 过素数，而 $a_p = 1 + p - N_p$ 是整数．由(12.1)式可知 $|a_p| \leqslant 2\sqrt{p}$，由此可知(12.2)式中右边的无穷乘积在 $s = \sigma + it$ 的实数部分 $\sigma > \frac{3}{2}$ 时是收敛的．所以在这个区域中定义出解析函数．$L(E,s)$称为椭圆曲线 E 的 L 函数．如果把 $s = 1$ 形式上代入到(12.2)式中，则有

$$L(E,1) = \prod_p \frac{1}{1 - a_p p^{-1} + p^{-1}}$$

$$= \prod_p \frac{p}{p - a_p + 1} = \prod_p \frac{p}{N_p}.$$

右边就是上述猜测中的无穷乘积．所以上述猜测可以说成：

若 $L(E,1) = 0$（即 $s = 1$ 为 $L(E,s)$的零点），则 $r \geqslant 1$．若 $L(E,1) \neq 0$，则 $r = 0$．

无穷乘积只在 $\sigma > \frac{3}{2}$ 时收敛，在 $s = 1$ 处并不能保证收敛，上述两位英国数学家猜想

$L(E,s)$可以解析延拓到 $s=1$ 处，并且猜想 $L(E,s)$在 $s=1$ 处的零点特性与 r 有更进一步的关系，我们简称它为 BSD 猜想．这个猜想是说

(1) 当 $\sigma>\frac{3}{2}$时由(12.2)式定义的复变函数可以解析开拓到整个复平面上的函数，并且没有奇点．并且有形如 $L(E,s)=f(s)L(E,2-s)$的函数方程．

(2) $L(E,s)$在 $s=1$ 处的零点阶数等于椭圆曲线 E 的有理点群 $E(\mathbb{Q})$的秩 r．

当 $L(E,1)\neq 0$ 时，还猜想非零实数 $L(E,1)$与椭圆曲线 E 的算术性质的关系，由于解释起来过于复杂，并且与费马猜想关系不大，此处从略．

到目前为止，人们只对一部分椭圆曲线证明了 BSD 猜想．1977 年，怀尔斯和他的老师科兹(John Coates)证明了：若 $L(E,1)\neq 0$，则 $r=0$，这在当时是一个突破．1985 年之后，又有更多的突破，但整个 BSD 猜想至今未能解决．

现在介绍关于椭圆曲线的另一个重要的猜想．看一下椭圆曲线 L 函数的级数(12.2)，并比较一下第 10 节赫克模形式的 L 函数欧拉无穷乘积展开式(10.13)($k=2$ 的情形)，它们在形式上完全一致．这导致如下的猜想

TSW 猜想 椭圆曲线的 L 函数 $L(E,s)$一定是对于某个群 $\Gamma_0(N)$的权 2 赫克模形式

的 L 函数.

这个猜想是从 1955 年到 1970 年逐渐形成的,目前对于此猜想的提出者也有一些争议. TSW 是日本数学家谷山丰(Taniyama)、志村五郎(Shimura)和法国数学家韦伊(Andre Weil)三个人的字头. 在 1955 年前后,谷山和志村就认识到椭圆曲线与权 2 赫克模形式的联系,并且计算了一些例子. 他们更多的是从代数几何的角度考虑问题. 1955 年 9 月在东京的一个国际学术讨论会上提出了类似的猜想. 在1967～1971年期间,韦伊对模形式和椭圆曲线作了深入的研



图 7　1993 年 6 月 23 日,怀尔斯在剑桥大学牛顿研究所演讲,宣布他对费马猜想的证明

究，提出了上述形式的猜想．由于角度不同，他们的叙述方式也有差异．从几何角度说成是：所有椭圆曲线都可以由模曲线来参量化．还有表示论的方式，说成：由椭圆曲线构作出的一种伽罗瓦群表示一定有某种特殊性质．人们逐渐认识到，这些表述方式彼此是同一个猜想的不同表述方式，我们在书中只挑选了其中比较容易讲的一种方式．这些不同的观点和角度是相互联系的．1993 年 6 月 23 日，怀尔斯在英国剑桥大学第一次宣布他证明了费马猜想的时候，他的演讲题目是：《模形式，椭圆曲线和伽罗瓦表示》．

TSW 猜想在数论中具有重要的位置，比如说，如果这个猜想成立，那么所有椭圆曲线的 L 函数都是某个权 2 赫克模形式的 L 函数．但是赫克已经证明了后者解析延拓成整个复平面上并且有函数方程．这表明由 TSW 猜想可推出 BSD 猜想的(1)．反过来，由于椭圆曲线 L 函数的系数有估计式 $|a_p| \leqslant 2\sqrt{p}$，所以那些对应的权 2 赫克模形式，其傅里叶系数也有类似地估计．所以，TSW 猜想使模形式与椭圆曲线之间建立了密切的关系．利用 TSW 猜想的其他几何或表示论的叙述形式，也可以推出关于椭圆曲线的许多深刻结果．1993 年 6 月，怀尔斯宣布他对于一类椭圆曲线证明了 TSW 猜想，并

由此推出费马猜想正确，引起极大的轰动．但是 TSW 猜想自身的意义不仅在于能证明费马猜想．到了 1994 年 8 月，尽管怀尔斯承认他的证明还不完善，由于他对于包括 BSD 猜想和 TSW 猜想在内的许多数论工作所作的贡献，仍旧邀请他在世界数学家大会的闭幕式上作了大会报告，并受到全场数学家的热烈欢迎．在费马猜想被证明之后，人们仍继续攻击 TSW 猜想．到了 2000 年，这个猜想已获完全解决．

TSW 猜想在数论中的地位，也可以从数学家的谈话中体现出来．怀尔斯的一位老师，英国著名数论学家科兹说："我在 1966 年开始从事研究工作，当时谷山和志村的猜想正席卷全世界．每个人都感到它很有意思，并开始认真地看待关于所有的椭圆曲线方程是否可以模形式化的问题．这是一段非常令人兴奋的时期．当然，惟一的问题是它很难取得进展．公正地说，虽然这个想法是漂亮的，但它似乎非常难以真正地证明，而这正是我们数学家主要感兴趣的一点．"

哈佛大学的梅祖尔(B. Mazur)教授是模形式和椭圆曲线理论的出色人物．他曾证明了所有椭圆曲线 E 的有限阶有理点群 $E(\mathbb{Q})_t$ 只有九种可能，这也是椭圆曲线理论中长期的一个猜想．他在评论 TSW 猜想时说："这是一个神

奇的猜想,推测每个椭圆曲线伴随一个模形式.但在一开始它就被忽视了,因为它太超前于它的时代.当它第一次被提出来时,没有被着手处理,因为它太使人吃惊.一方面是椭圆曲线,另一方面是模形式,这两个数学分支都被集中地但各自独立地研究过.研究椭圆曲线的数学家可能并不精通模形式,反过来也是一样.于是,谷山-志村猜想出现了.这个重大的猜想说,在这两个完全不同的世界之间存在一座桥.数学家喜欢建造桥梁."

是的,数学家喜欢建造桥梁.TSW 猜想在模形式和椭圆曲线之间建了一座桥,使人感到吃惊.几十年之后,又有人在 TSW 猜想和费马猜想之间建了一座桥,更使人感到振奋.这件事发生在德国一座风景优雅而安静的小城中.

13 怀尔斯面壁8年

德国黑森林州中部有一个小城，名叫奥伯沃尔法赫(Oberwerfach).世界上许多数学家都知道这里有一处风景优雅而恬静的地方，每年都安排50余个数学讨论会.上个会议的数学家于星期六离开，而下一个会议的参加者于星期日到达.室内有图书，期刊，黑板，投影仪和咖啡，没有电视和尘世的喧闹.1984年秋天，一群优秀的数论学家在这里聚会，讨论椭圆曲线方面一些最新结果.来自德国的符莱(G.Frey)走上讲台，先写下了费马方程

$$x^p + y^p = z^p \quad (p \geqslant 3),$$

然后转过身来对听众说，"如果此方程有正整数解$(x, y, z) = (A, B, C)$"，接着又转过身去写出一个椭圆曲线

$$y^2 = x(x + A^p)(C^p - x).$$

他接着说，他研究了这条椭圆曲线，发觉它非常古怪，古怪到这条曲线不可能被模形式化.这相

当于说，TSW 猜想对这条椭圆曲线是不成立的！这又等于说：若费马猜想不成立，则可构造出一条椭圆曲线使 TSW 猜想不成立．或者反过来说：由 TSW 猜想成立可以推出费马猜想成立．

符莱在讨论会上只是讲述了他的思路，论述中有缺欠．专家们也看出缺欠，但在自由交流思想的讨论中这是不足为奇的，重要的是新奇的想法．演讲过后，听众纷纷拿走符莱演讲的预印本，回到自己的研究所，试图弥补符莱的缺欠．

美国加州大学柏克莱分校的教授里贝特（K. Ribet）也是符莱演讲的听众之一．这位数论专家回去之后也迷恋于研究为什么符莱曲线不能模形式化．但是 18 个月之后，他和所有其他人一样没有做出任何结果．1986 年夏天，哈佛大学的梅祖尔访问柏克莱并出席世界数学家大会，在咖啡店遇到里贝特．在一阵闲谈之后，里贝特向梅祖尔讲述他在这 18 个月里对符莱曲线所作的考虑，“我提到我已经证明了非常特殊的情形，但是我不知道下一步如何推广以得到整个证明．”

梅祖尔一边喝咖啡一边思考着里贝特的讲述．突然间他停了一下，注视着里贝特：“难道你不明白？你已经完成了它！你需要的只是加上

一些模结构，然后再做一遍你的论证就行了．”

里贝特后来回忆说：“我高兴得像上了天似地回到住所，满脑子想着：天哪！这难道真是对的吗？我完全被迷住了．……大约一个或两个小时，我已经写完了一切，确信我已掌握了关键的步骤．……世界数学家大会当然有成千的数学家参加，我有点随便地对几个人说我已经证明了由 TSW 猜想可推出费马猜想．这消息像野火般传开来，他们问我：你已经证明了符莱椭圆曲线不能模形式化吗？……我叫道：是的，我已经证明了．”

现在，符莱的想法得到确认，如果想证明费马猜想，就要去攻击 TSW 猜想．后一件事连里贝特也持悲观态度：“绝大多数人相信谷山－志村猜想是完全无法接近的，我是其中的一个…….怀尔斯大概是地球上敢大胆梦想可以证明这个猜想的极少数人之一．”

怀尔斯没有出席那次世界数学家大会，但是他后来回忆会后不久发生的事：“那是 1986 年夏末的一个傍晚，当时我在朋友家中喝冰茶．谈话间他随意告诉我，里贝特已经证明了谷山-志村猜想和费马猜想之间的联系．我感到极大的震动．我记得那个改变我生命历程的时刻，因为这意味着为了证明费马猜想，我必须做的一切就是去证明谷山-志村猜想．它意味着我童年

的梦想现在成了体面的值得去做的事．我懂得我决不能让它溜走．我十分清楚我应当回家去研究谷山－志村猜想．"



图8　安德鲁·怀尔斯

安德鲁·怀尔斯于 1953 年生于英国剑桥．

1963 年，10 岁的怀尔斯已经对数学着迷，喜欢做题目．从学校回家的路上，常到地区图书馆看书．在这里他看到一本介绍费马猜想的书．30 年后他提到当时的感受："它看上去如此简单，但历史上所有大数学家都未能解决它．这里正摆着我——一个 10 岁的孩子——能理解的问题．从那时刻起，我知道我永远不放弃它，我必须解决它．"这就是他所说的"童年的梦想．"

1974 年，怀尔斯在剑桥大学莫尔顿学院毕业．1977 年在克莱尔学院获博士学位．这时，他与他的老师科兹发表了关于 BSD 猜想的突破性结果(如果 $L(E,1)\neq 0$，则椭圆曲线 E 的秩为 0)．1977～1980 年，他在美国哈佛大学做助理教授．1981 年在普林斯顿高等研究院任研究员．这期间，他与哈佛大学的梅祖尔合作证明了有理数域上的岩泽健吉主猜想(这是分圆域理论的一个重要猜想)．1982 年任普林斯顿大学教授．1984 年至今任该校讲座教授．从 1988 年到 1990 年他还兼任牛津大学皇家协会研究教授．

从 24 岁开始，怀尔斯在世界上最有声誉的这些大学和研究所工作，是基于他对数论研究的杰出贡献．特别在模形式、分圆域理论和椭圆曲线方面，他是数学界公认的最优秀专家之一．所以当他在 1986 年得知里贝特确认由 TSW

猜想可以推出费马猜想之后,便放弃了所有与此无关的研究工作,也不再过多地出席学术会议.除了在普林斯顿大学教书之外,一头扎进他自己的顶楼书房里,潜心于思考 TWS 猜想.



图 9　安德鲁·怀尔斯在 1986 年意识到有可能通过谷山-志村-韦伊猜想证明费马猜想

从一开始，怀尔斯就作了一个重大的决定：要完全独立和保密地进行研究．面壁 8 年之后他解释自己的这种动机是希望自己的工作不受干扰．“我意识到与费马猜想有关的任何事情都会引起太多人的兴趣．你确实不可能很多年都使自己精力集中，除非你的专心不被他人分散，而当旁观者太多时是不可能作到的”．在以后的几年里，他一直守口如瓶．对于自己取得的进展从不与别人讨论或发表．科兹回忆起这段时期与怀尔斯的交往：“我记得在许多场合对他讲：与费马猜想的这种联系确实非常好，但要想证明谷山－志村猜想仍然是毫无希望的．他当时只是对我笑笑．”里贝特也不知道怀尔斯暗中进行的工作：“这大概是我知道的惟一例子，一个人进行了这么长时间的研究而不公开他在做什么，也不谈论自己的进展．在我的经历中这是前所未闻的．”惟一知道怀尔斯秘密的人是他的妻子内达（Nada）．“我的妻子是惟一知道我一直在从事费马猜想研究的人．度蜜月时我告诉了她，那时我们才结婚几天．她听说过费马猜想，不过在那个时候她一点也不知道它对于数学家所具有的传奇式的意义．”

在这段时间里，“基本上整段时间里围绕在我脑海中的是这件事．早晨醒来想到的第一件事就是它，然后一整天在思考它，在梦中我也会

思考它.只要没有分心的事,我会整天翻来复去想这件事."“我有时候在纸上潦草地写上几笔,或者说是乱涂.它们不是什么要紧的东西,只是下意识地乱涂乱写.我从不用计算机."经过一年多的摸索,他认定采取用伽罗瓦表示的途径攻击 TSW 猜想,但是在 1988 年 3 月 8 日,华盛顿邮报和纽约时报登在第一版的一条消息使他大吃一惊.

1988 年 3 月 7 日,日本几何学家宫岗洋一在联邦德国首都波恩的马普数学所宣布他证明了费马猜想.从 1983 年法廷斯证明莫德尔猜想之后,人们探索解决费马猜想的几何途径(见第 7 节),愈来愈多的几何高手对费马猜想发生兴趣.宫岗是一位优秀的年轻几何学家,他在马普数学所的演讲中宣布他证明了微分几何中的一个不等式,并且由这个不等式可推出费马猜想.过了两个星期,宫岗公布了关于这个不等式的 5 页纸的证明.世界各地的几何学家和数论学家逐行地检查他的证明.几天之内就有人察觉到证明中有漏洞.又过了两个星期,法廷斯更明确指出宫岗证明中的错误所在.一批数论学家试图帮助宫岗补救错误,历时两个月也未能成功.人们逐渐感到补救的难度不亚于费马猜想本身.不久,报界又刊登了一个更正,说这个 300 多年的谜仍未解决.这件事的余波传到纽

约第8街的地铁车站里．有人仿照费马在丢番图《算术》一书的页边上所写的评注，在墙上涂写了打油诗：

$x^n + y^n = z^n$ 没有解

对此，我已经发现一个真正美妙的证明

可惜我现在没有时间写下它

因为我的火车正在开来

怀尔斯终于松了一口气，但是他的工作进展不大．这期间他当了两次父亲，"我放松一下情绪的惟一方式是与孩子们在一起．年幼的孩子们当然对费马猜想毫无兴趣，他们只需要听故事，不想让你做任何其他事情．"到了1991年夏天，在普林斯顿当了5年的隐士之后，他去波士顿出席关于椭圆曲线的一个高水平会议．怀尔斯受到朋友们的热情欢迎，人们很高兴又见到他．在会上，科兹告诉他有一位学生弗拉赫(M. Flach)正在用前苏联数学家柯里瓦金(Kolyvagin)的方法研究椭圆曲线．回到普林斯顿之后，怀尔斯花了几个月的时间去理解柯里瓦金和弗拉赫方法．在取得一些进展之后，他愈来愈相信这种方法对于证明TSW猜想是有效的．

到了1993年1月，怀尔斯已经完全相信用此方法可以证明TSW猜想．这个时候，他感到需要有人帮助了．"那一年我工作得异常努力，

试图使柯里瓦金和弗拉赫方法能成功，但是这个方法我并不真正熟悉．其中有许多很难的代数，需要我学习许多新的数学．于是，大约在1993年1月份的上半月，我决定有必要向一个人吐露秘密，而他应该是一位我正在使用的那一类几何方法方面的专家．我需要非常小心地挑选这个我要告知秘密的人，因为他必须保守住秘密．我选择了尼克·凯兹．”

尼克·凯兹(Nick Katz)是怀尔斯在普林斯顿大学数学系的同事，资深的代数几何和数论专家．他回忆起当时的情景：“有一天怀尔斯在喝茶休息时走到我身边，问我是否能一起到他的办公室去，他有些事想和我谈．我一点也不知道他想和我谈什么．我和他一起到了他的办公室，他关上了门．然后说他认为他可以证明谷山－志村猜想．我大吃一惊，目瞪口呆，这真是异想天开．

他解释说证明中有一大部分是依靠他对弗拉赫和柯里瓦金的工作所作的扩展，但是非常专门．他对证明中这非常专门性的部分感到没有把握，想和某个人一起讨论这部分，以保证它是正确的．他认为我是帮助他核对的正确人选．但是我认为，他所以选中我还有别的原因：他相信我会守口如瓶，不会告诉别人．

在6年的孤军作战之后，怀尔斯终于向别

人吐露了秘密.他们两人感到工作量很大,需要每周定时讨论,但又不能被人注意,因为凯兹是位朋友众多,交际广泛的人物.最后两人商定的办法是由怀尔斯为研究生开设课程,名为"椭圆曲线的计算",而凯兹成为听众之一(教授听教授的课这在美国是习以为常的).这个课程的名称不会引起任何人的注意.课程一开始,怀尔斯就讲述使研究生莫名其妙的演算.学生们一个个地走掉,几个星期之后,听众只剩下凯兹一人.

凯兹仔细地听着怀尔斯演算的每一步.在课程结束时,凯兹认为怀尔斯的方法似乎是完全可行的.而怀尔斯也努力完成证明.

"5月末的一个早晨,内达和孩子们一起出去了,我坐在书桌旁思考着这剩下的一族椭圆方程式.我随意地看一下梅祖尔的一篇论文,恰好其中有一句话引起我的注意.它提到19世纪的一个构造.我突然意识到这种结构可以使我对这最后一族椭圆曲线采用柯里瓦金-弗拉赫方法,我一直工作到下午,忘记了吃午饭.到了三四点钟的时候,我真正确信这将解决剩下的问题."

经过7年的潜心研究,怀尔斯认为已经到了向全世界公布的时候."这样,到了1993年5月,我确信我已掌握了整个费马猜想.恰好在6

月份剑桥有一个会议，我想这也许是我宣布这个证明的好地方，它是我古老的家乡，我曾是那里的研究生.”

在剑桥大学牛顿研究所举行的会议，名为“L函数和算术”．组织者正是怀尔斯作研究生的导师科兹、梅祖尔、里贝特、柯里瓦金、怀尔斯过去的学生鲁宾（K. Rubin）……数论的世界顶尖高手陆续到来．这个时候，已经在传说怀尔斯证明了费马猜想．所以在怀尔斯演讲之前，科兹问他：“你究竟证明了什么？我们要不要告诉新闻界？”怀尔斯只是微微地摇了一下头，依然紧闭双唇．“他确实在为高度戏剧性的场面作准备.”

怀尔斯的演讲分三次进行，题目为“模形式、椭圆曲线和伽罗瓦表示”．6月21日的第一次演讲是一般性的．演讲结束后，鲁宾的电子邮件飞向全世界各地：

日期：1993年6月21日，星期一，13点33分6秒

标题：怀尔斯

各位，安德鲁今天作了他的第一次报告．他没有宣布对谷山－志村猜想的证明，但是他正在向那个方向迈进．他还有两次报告．关于最后的结果他仍然非常保密．……

卡尔·鲁宾.

到了第二天，听众增加许多.怀尔斯讲了过渡性演算.这些演算表明他指向谷山－志村猜想的意图，但听众仍不清楚他是否足以能证明它从而征服费马猜想.下面是鲁宾在第二天发出的电子邮件

“今天的报告中无更多实质性内容.他叙述了我昨天猜到的方向上关于提升伽罗瓦表示的一般性定理.它似乎并不适用于所有的椭圆曲线.但精妙之处将出现在明天.

我真的不知道为什么他要以这种方式来进行演讲.显然他自己知道明天讲什么，这是他10多年来一直从事的规模非常宏大的工作.他似乎对此很自信.

我会告诉你们明天的情况.”

到了第三天，剑桥数学界的每个人都来听讲.运气好的人挤进了演讲厅，其他人只能在走廊里踮起脚，透过窗子向里凝视.里贝特回忆：“我到得比较早，和梅祖尔一起坐在前排.我带着照相机以便记录这个重大事件.当时的气氛充满了激情，人们非常兴奋.大家肯定都意识到我们正参与一个历史性事件.”梅祖尔说：“我从未见过如此辉煌的演讲，充满了如此奇妙的思想，具有如此戏剧性的紧张，准备得如此之好.”怀尔斯本人回忆当时的情景：“虽然新闻界已听到有关演讲的风声，很幸运他们没有来听讲.但

是听众中许多人拍摄了演讲结束时的镜头.研究所所长事先准备了一瓶香槟酒.当我宣读证明时,会场上保持特别庄重的寂静.然后当我写完费马猜想的命题,并且说:我想我就在这里结束.接着会场上爆发出一阵持久的鼓掌声."里贝特对当时场景的形容:"人们彼此对望着,喊道:我的天啊!要知道我们刚才亲眼目睹了一个多么伟大的事件!……下一位报告人是一个名叫里贝特的人,那就是鄙人.我作了演讲,人们作了笔记,也鼓了掌.可是在场的每一个人,包括我自己,对我在演讲中讲了些什么都没有丝毫的印象."

鲁宾在第三天发出的电子邮件向全世界热情而详细地介绍了怀尔斯证明思想:怀尔斯对于所有"半稳定"的椭圆曲线证明了 TSW 猜想.由于符莱曲线是半稳定的,这就推出了费马猜想是对的.第二天,电视台和科学新闻记者大批来到牛顿数学所,要求采访"本世纪最杰出的数学家".《卫报》说:"数论因数学的最后之谜而看涨".《世界报》的头版标题为"费马猜想得到解决".《人物》(People)杂志把他和戴安娜王妃等人一起列为本年度 25 位最具魅力的人物之一.一家国际制衣大企业请这位温文尔雅的天才为他们的新系列男装作广告.

我们在第 7 节说过,1816 年,巴黎科学院

曾为费马猜想设立了3000法朗的奖金(后来因法朗贬值又第二次设奖).1908年,德国又以10万马克授予第一个证明费马猜想的人.并且有严格规定:

(1)……只考虑在定期刊物上以专著形式发表或在书店中出售的数学专题论著……

…………

(9)如果到2007年9月13日尚未颁布此奖,将不再继续接受此奖.

就在媒体热情报道的同时,有资格的数学家们却冷静地进行审查工作.怀尔斯将他的200页手稿投寄《数学发明》(Inventiones Mathematicae)杂志.该杂志的数论编辑梅祖尔破例指定了六个审稿人,每人负责其中的一章.凯兹被分到审查第3章,他邀请巴黎的伊卢齐(L. Illusie)合作审稿."我们只是逐行审阅原稿,想办法确保不存在错误."每天都要通过电子邮件和怀尔斯讨论不清楚的地方.8月23日发现一个问题,怀尔斯认为并不严重.但是凯兹持执著的态度.到了9月份凯兹感到这是个重大的缺欠.这给两个人都同时带来巨大的压力,但是他们对外闭口不言.另一个审稿人里贝特被人称为"费马信息咨询所".被邀请到处去作报告.几个月后,人们询问他:论文怎么样了?为什么没有进一步消息?数学界之间的电子邮件也出现

了"怀尔斯证明中有缺欠吗?","关于费马漏洞"这样的标题.外界对他的证明迟迟不公开而产生的情绪不断增长.到了1993年12月4日,怀尔斯认识到他不能永远保持沉默,发出了电子邮件:

"由于存在着对我的关于谷山-志村猜想和费马猜想的工作情况有种种推测,我将对情形作一简短说明.在检验过程中发现许多问题,大部分已经解决,但是有一个特别的问题我还没有解决……我相信在不远的将来能够用我在剑桥演讲中解释的想法完成它."

这个时期,怀尔斯陷入了苦闷状态,他向普林斯顿大学数学系的另一个朋友萨纳克(P. Sarnak)承认情况已面临绝境,他准备承认失败.萨纳克向他暗示:困难的一部分原因是由于怀尔斯缺少一个可以信赖的进行日常讨论的人.并建议他寻找一个朋友,再试一次.怀尔斯认真考虑了这个建议,请剑桥大学讲师泰勒(R. Taylor)来普林斯顿与他一起工作.泰勒是怀尔斯文章的审稿人之一,也是怀尔斯以前的学生.他们工作了半年多之后,仍没有大的突破.

1994年8月,世界数学家大会在瑞士的苏黎士举行.大会每四年举行一次,并且要为不超过40岁的数学家颁发世界数学最高荣誉:菲尔

兹奖.怀尔斯对费马猜想的工作没有完全确认,而且那时他已经超过了40岁.但是由于他对数论的其他重要贡献,仍被邀请在大会闭幕式上作了一小时报告.

8月底,怀尔斯对泰勒说,他看不出继续努力会有什么指望.泰勒建议再坚持一个月.如果到9月底还修改不好,他准备回剑桥.这时怀尔斯的心态比较平静.尽管费马猜想的证明没有完成,他的全部工作仍是很有价值的,他所提供的一大套新的技术和策略,可得到其他结果.他已感到失败不是羞耻,开始适应受到打击后的境遇.他准备到9月底公开承认他们的失败并发表有缺欠的证明,希望其他人有机会去研究它.当泰勒试图探索别的方法时,怀尔斯决定最后一次审视一下他失败的原因.

"9月19日早晨,我坐在桌旁检查柯里瓦金和弗拉赫方法.这倒不是因为我相信自己能使它行得通,而是我认为至少能解释为什么它行不通,……突然间,完全出乎意料,我有了一个难以置信的发现.…….单靠岩泽理论不足以解决问题,单靠柯里瓦金和布拉赫方法也不足以解决问题,它们结合在一起却可以完美地相互补充.""……我无法理解我怎么会没有发现它,足足有20分钟我呆望着它不敢相信.然后到系里转了一圈,又回到桌旁.情况确实如此.

我无法控制自己，我太兴奋了．这是我工作经历中最重要的时刻，我所做的工作中再也没有哪一件会具有这么重要的意义．……这是我感到轻松的第一个晚上．我把事情放到第二天再去做．第二天早晨我又一次作了核对，到 11 点钟时我完全放心了．"

1994 年 10 月 25 日，卡尔·鲁宾再次向数学界发了电子邮件：

标题：费马猜想的最新情况

到目前为止，2 份手稿已经寄出：

《模椭圆曲线和费马最后定理》作者：怀尔斯

《某些黑克代数的环论性质》作者：泰勒和怀尔斯．

第一篇论文（长），除了别的结论外，宣布了费马猜想的一个证明，它的关键一步依赖于第二篇论文（短）．

…………

整个论证的概要与怀尔斯在剑桥描述的相似．由于不再用欧拉系，新的处理方法比原来大大简单和快捷．（事实上，在看到这些手稿之后，法廷斯已对那部分推导提供了进一步的重大简化．）……

卡尔·鲁宾（俄亥俄州立大学）

两篇论文共 130 页，是历史上核查得最彻

底的数学稿件，最终发表在《数学年刊》(Annals of Mathematics)上(1995 年 5 月). 怀尔斯的名字再次出现在《纽约时报》的头版上，不过这一次的标题《数学家称经典之谜已解决》和另一则报道《宇宙年龄的发现提出新的宇宙之谜》比较，有点相形见绌. 尽管记者们的热情稍有减退，但数学家们并未忽视这个证明的巨大意义. 科兹认为"这个最终的证明可与原子分裂或发现 DNA 的结构相比，是人类智力活动的一曲凯歌."里贝特的说法是："我想假如有人被遗弃在一个无人的荒岛上，而他只带着这篇论文，那么他会有大量的精神食粮."1996 年 3 月，怀尔斯和朗兰兹(Langlands)分享了 10 万美元的沃尔夫奖. 这是和菲尔兹奖齐名的数学奖，但通常授予一生中对数学有杰出贡献的资深数学家，这个奖给一位四十多岁的年轻人尚属首次. 沃尔夫奖委员会认为，怀尔斯的证明就其本身来说是一个使人震惊的成就，而同时它也给雄心勃勃的朗兰兹纲领注入了生命力. 1997 年 6 月 27 日，怀尔斯收到德国为费马猜想设立的价值 5 万美元的奖金. 1998 年 8 月，世界数学家大会在柏林举行. 由于怀尔斯已超过菲尔兹奖的 40 岁年限，大会开幕式上破天荒地为怀尔斯颁发了一项"特别贡献奖"：国际数学联盟银牌(被数论学家查杰尔(D. Zagier)戏称为"量子化的菲

尔兹奖")．他得到了比会上宣布的菲尔兹奖新得主们更为热烈的掌声．第二天他在大会上作了题为《数论 20 年》的演讲．

到此为止，再对费马猜想说更多的话似乎已属多余．只想对于和怀尔斯同时获沃尔夫奖的朗兰兹再写上几笔．朗兰兹纲领是一个宏大的数学规划、看法和哲学思想．它所涉及的领域不仅是数论，而是几何、代数和分析诸方面的融合和统一．简言之，朗兰兹纲领给出对数学世界的一种观点：如果我们研究某个数论、代数或几何的对象 S，并且我们试图寻找某种解析函数（称为 S 的 ξ 函数或 L 函数），使此函数的解析特性能反映 S 的几何、数论或代数特性，那么这个函数应当来源于与 S 有关的某个群的某种自守表示，即它应当对应于模形式的一种推广：自守形式．如果详细解释朗兰兹纲领，那就是另外一个数学故事了．

附　　录

关于费马猜想的书籍和介绍文章有许多种.这里推荐最近出版的两本中文书籍.

[1] 胡作玄.350年历程:从费马到维尔斯.山东教育出版社.1996年

[2] Simon Singh著.薛密译.费马大定理.上海译文出版社.1998年